

```

;
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
;
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 11
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 12
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
```

```

US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 13
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (Blys)
; FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 14
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
```

Mon Apr 19 13:27:24 2004

```

; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
DB      478 TESLVNRRP 486

RESULT 15
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 45; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
DB      478 TESLVNRRP 486

Search completed: April 19, 2004, 12:55:00
Job time : 4.78947 secs
```





MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-n)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"  
US-08-153-799-14

Query Match 100.0%; Score 51; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.065;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHEAKRMP 9  
|||||  
DB 439 KHEAKRMP 447

## RESULT 2

US-08-448-196A-3  
Sequence 3, Application US/08448196A  
Patent No. 5780594  
GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C.  
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
TITLE OF INVENTION: RELATED PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NASA  
STREET: MARSHALL SPACE FLIGHT CENTER  
CITY: HUNTSVILLE  
STATE: ALABAMA  
COUNTRY: USA  
ZIP: 35812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448.196A  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BROAD JR., ROBERT L.  
REGISTRATION NUMBER: 18,757  
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
TELEPHONE: 205-544-0021  
TELEFAX: 205-544-0258  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal

US-08-448-196A-3  
Query Match 100.0%; Score 51; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.065;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHEAKRMP 9  
|||||  
DB 439 KHEAKRMP 447

## RESULT 3

US-08-984-176-1  
Sequence 1, Application US/08984176  
Patent No. 5948609  
GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C  
APPLICANT: HO, JOSEPH X  
APPLICANT: RUKER, FLORIAN  
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER  
FILE REFERENCE: 08/984,176  
CURRENT APPLICATION NUMBER: US/08/984,176  
CURRENT FILING DATE: 1997-12-03  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 1  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-984-176-1

Query Match 100.0%; Score 51; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.065;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHEAKRMP 9  
|||||  
DB 439 KHEAKRMP 447

## RESULT 4

US-08-702-572-2  
Sequence 2, Application US/08702572  
Patent No. 5965386  
GENERAL INFORMATION:  
APPLICANT: Kerry-Williams, Sean M  
APPLICANT: Gilbert, Sarah C  
TITLE OF INVENTION: Yeast Strains and Modified Albumins  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Centeon L.L.C.  
STREET: 1020 First Avenue  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406-1310

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,572  
FILING DATE: 11-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 95/23857  
FILING DATE: 1-MAR-1995  
APPLICATION NUMBER: GB 9404270.2  
FILING DATE: 5-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Naomi Biswas  
REGISTRATION NUMBER: 38,384  
REFERENCE/DOCKET NUMBER: CE0114 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610/878/4294  
TELEFAX: 610/878/4221  
INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2
Query Match 100.0%; Score 51; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cartoll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2
Query Match 100.0%; Score 51; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match 100.0%; Score 51; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3
Query Match 100.0%; Score 51; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```

APPLICANT: Bair, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschoep, Juerg F.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DiGiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9108Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 51; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
Db 463 KHPEAKRMP 471

RESULT 9  
US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CFA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 51; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
Db 463 KHPEAKRMP 471

RESULT 10  
US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 51; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
Db 463 KHPEAKRMP 471

RESULT 11  
US-09-976-594-977  
; Sequence 977, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 977  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1  
US-09-976-594-977

Query Match 100.0%; Score 51; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
Db 463 KHPEAKRMP 471

RESULT 12  
PCT-US95-04075-3  
; Sequence 3, Application PC/TUS9504075  
; GENERAL INFORMATION:  
; APPLICANT: AMGEN INC.

;; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
;; TITLE OF INVENTION: Protein  
;; NUMBER OF SEQUENCES: 33  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Amgen Center, Patent Operations/RRC  
;; STREET: 1840 DeHavilland Drive  
;; CITY: Thousand Oaks  
;; STATE: California  
;; COUNTRY: U.S.  
;; ZIP: 91320-1789  
;; COMPUTER READABLE FORM: disk  
;; MEDIUM TYPE: Floppy  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/04075  
;; FILING DATE:  
;; CLASSIFICATION:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 609 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 100.0%; Score 51; DB 5; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
|||  
Db 463 KHPEAKRMP 471

RESULT 13  
US-08-797-689-2  
; Sequence 2, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittion, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US95-04075-3

;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38 619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-797-689-2

Query Match 100.0%; Score 51; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9  
|||  
Db 463 KHPEAKRMP 471

RESULT 14  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. 6686179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittion, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38 619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:

```

;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 610 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 51; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. NO. 0.069;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHPEAKRMP 9
Db      463 KHPEAKRMP 471

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
;   APPLICANT: Bell et al.
;   TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
;   FILE REFERENCE: PF556
;   CURRENT APPLICATION NUMBER: US/10/153,064
;   CURRENT FILING DATE: 2002-05-24
;   PRIOR APPLICATION NUMBER: 60/293,212
;   PRIOR FILING DATE: 2001-05-25
;   NUMBER OF SEQ ID NOS: 137
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 133
;   LENGTH: 651
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 51; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. NO. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHPEAKRMP 9
Db      505 KHPEAKRMP 513

Search completed: April 19, 2004, 12:05:21
Job time : 2.3795 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.70637 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_462\_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.78.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 71    | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 2          | 67    | 94.4        | 608    | 2 S57632 | serum albumin prec |
| 3          | 64    | 90.1        | 600    | 2 A47391 | serum albumin prec |
| 4          | 64    | 90.1        | 605    | 1 ABPGS  | serum albumin prec |
| 5          | 64    | 90.1        | 607    | 1 ABBOB  | serum albumin prec |
| 6          | 64    | 90.1        | 607    | 1 ABSHS  | serum albumin prec |
| 7          | 64    | 90.1        | 608    | 1 ABRTS  | serum albumin prec |
| 8          | 63    | 88.7        | 607    | 1 ABHOS  | serum albumin prec |
| 9          | 60    | 84.5        | 609    | 2 UC5838 | albumin - Mongolia |
| 10         | 59    | 83.1        | 453    | 2 A05139 | serum albumin - mo |
| 11         | 43    | 60.6        | 553    | 2 D83640 | hypothetical prote |
| 12         | 43    | 60.6        | 733    | 2 E71937 | flagellar biosynth |
| 13         | 43    | 60.6        | 733    | 2 A64650 | flagellar biosynth |
| 14         | 40    | 56.3        | 342    | 2 S63404 | hypothetical prote |
| 15         | 40    | 56.3        | 604    | 2 B3C289 | conserved hypochet |
| 16         | 40    | 56.3        | 640    | 2 T29784 | hypothetical prote |
| 17         | 39    | 54.9        | 577    | 2 T45333 | hypothetical prote |
| 18         | 39    | 54.9        | 728    | 2 B84415 | translation elonga |
| 19         | 39    | 54.9        | 728    | 1 S07558 | probable DNA-bind  |
| 20         | 39    | 54.9        | 753    | 2 G87178 | methane monooxygen |
| 21         | 38    | 53.5        | 169    | 2 C39049 | hypothetical prote |
| 22         | 38    | 53.5        | 172    | 2 T70789 | hypothetical prote |
| 23         | 38    | 53.5        | 309    | 2 G83044 | lipase lipC PA4813 |
| 24         | 38    | 53.5        | 457    | 2 T02786 | serine-tRNA ligase |
| 25         | 38    | 53.5        | 539    | 2 T02705 | hypothetical prote |
| 26         | 38    | 53.5        | 690    | 2 T97164 | flagellar biosynth |
| 27         | 37    | 52.1        | 277    | 2 T36312 | hypothetical prote |
| 28         | 37    | 52.1        | 277    | 2 AB3642 | protein ybis precu |
| 29         | 37    | 52.1        | 383    | 2 T50775 | probable translati |

hemolin - cecropia  
hemolin precursor  
neutral metallopro  
hypothetical prote  
hypothetical prote  
hypothetical Myb f  
pled protein - Syn  
probable translati  
translation initia  
probable translati  
phosphoribosylamin  
filamentous hemagg  
hypothetical prote  
hypothetical prote  
lmbN protein - Str  
DNA-directed DNA p

30 37 52.1 413 2 S65948  
31 37 52.1 413 2 A37778  
32 37 52.1 438 2 S42581  
33 37 52.1 477 2 T16546  
34 37 52.1 482 2 G96616  
35 37 52.1 661 2 T38176  
36 37 52.1 829 2 S75776  
37 37 52.1 900 2 T47732  
38 37 52.1 900 2 T50773  
39 37 52.1 935 2 T50774  
40 37 52.1 1353 1 AUFFEM  
41 37 52.1 4152 2 T31102  
42 37 52.1 4919 2 T31105  
43 36 50.7 113 2 A81439  
44 36 50.7 275 2 S44961  
45 36 50.7 331 2 T09748

#### ALIGNMENTS

##### RESULT 1

ASHUS  
serum albumin precursor [validated] - human  
N:Alternate names: preproalbumin  
N:Contains: kinetensin  
C:Species: Homo sapiens (man)  
C:Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000  
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S31  
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu:  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col:  
A:Reference number: A93743; MUID:82081882; PMID:6171778  
A:Accession: A93743  
A:Molecule type: mRNA  
A:Residues: 1-419, 'K', 421-609 <LAW>  
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2:  
R:Dugaiczyk, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A:Reference number: A93936; MUID:82105994; PMID:6275391  
A:Accession: A93936  
A:Molecule type: mRNA  
A:Residues: 1-120, 'G', 122-609 <DUG>  
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
R:Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, I.  
J. Biol. Chem. 261, 3244-3251, 1986  
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A:Reference number: I39427; MUID:86140099; PMID:2419329  
A:Accession: I39427  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <URA>  
A:Cross-references: GB:IMJ3075; NID:G178330; PIDN:AAA51688.1; PID:G553173  
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A:Reference number: I59286; MUID:94181575; PMID:8134387  
A:Accession: I59286  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 282-290, 'KSRPDLQ' <WAT>  
A:Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033  
A>Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A:Reference number: I59313; MUID:94294404; PMID:8022807  
A:Accession: I59313  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 589-590, 'ALPRVKNLLQVLP' <MAD>  
A:Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
 R;Menaya, J.; Parvillia, R.; Ayuso, M.S.  
 submitted to the EMBL Data Library, March 1995  
 A;Reference number: G08292  
 A;Accession: G01747  
 A;Status: translated from GS/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-120, 'G', 122-455 <MEN>  
 A;Cross-references: EMBL:U22961; NID:G763428; PID:G763431  
 R;Ledgerwood, B.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
 Biochem. J. 308, 321-325, 1995  
 A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
 A;Reference number: S55314; MUID:95275251; PMID:7755581  
 A;Accession: S55314  
 A;Molecule type: protein  
 A;Residues: 19-27 <LED>  
 R;Meloun, B.; Moravek, L.; Kostka, V.  
 FEBS Lett. 58, 134-137, 1975  
 A;Title: Complete amino acid sequence of human serum albumin.  
 A;Reference number: A91420; MUID:76187907; PMID:1225573  
 A;Accession: A91420  
 A;Molecule type: protein  
 A;Residues: 25-117, 'EO', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
 R;Roehr, U.; Spittler, G.; Tripiet, D.  
 Justus Liebig's Ann. Chem. 9, 881-884, 1988  
 A;Title: Isolation and structure elucidation of middle-molecular weight peptides from un  
 A;Reference number: S06422  
 A;Note: this paper is in German, with an English abstract  
 A;Accession: S06422  
 A;Molecule type: protein  
 A;Residues: 25-48 <RO>  
 R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
 Arch. Biochem. Biophys. 305, 593-599, 1993  
 A;Title: Mass spectrometric identification of modifications to human serum albumin treat  
 A;Reference number: S36882; MUID:93384321; PMID:8373198  
 A;Accession: S36882  
 A;Molecule type: protein  
 A;Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>  
 R;Kausler, E.; Spittler, G.  
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
 A;Title: Bruchstuecke aus Albumin und Beta(2)-Mikroglobulin - Bestandteile der Mittelmol  
 A;Reference number: S17599; MUID:92126241; PMID:1772598  
 A;Accession: S17599  
 A;Molecule type: protein  
 A;Residues: 25-54, 354-357, 431-447 <KAU>  
 A;Note: 49-Leu was also found  
 R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
 J. Immunol. 143, 1680-1684, 1989  
 A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
 A;Reference number: A45800; MUID:89341406; PMID:2474609  
 A;Accession: A45800  
 A;Molecule type: protein  
 A;Residues: 166-173 <CAR>  
 R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
 A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
 A;Reference number: A03239; MUID:86242180; PMID:3087352  
 A;Accession: A03239  
 A;Molecule type: protein  
 A;Residues: 166-173, 'L', <MOG>  
 R;Galliano, M.; Minichioti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
 A;Title: Mutations in genetic variants of human serum albumin found in Italy.  
 A;Reference number: A38255; MUID:91062352; PMID:2247440  
 A;Accession: C38255  
 A;Molecule type: protein  
 A;Residues: 76-111 <GAL1>  
 A;Accession: B38255  
 A;Molecule type: protein  
 A;Residues: 82-105, 'K', 107-110 <GAL2>  
 A;Note: this variant is designated albumin Vibo Valentia  
 A;Accession: A38255  
 A;Molecule type: protein

A;Residues: 76-83, 'K', 85-106 <GAL3>  
 A;Note: this variant is designated albumin Torino  
 R;Minichioti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
 Eur. J. Biochem. 214, 437-444, 1993  
 A;Title: The structural characterization and bilirubin-binding properties of albumin He  
 A;Reference number: S33298; MUID:93292504; PMID:8513793  
 A;Accession: S33298  
 A;Molecule type: protein  
 A;Residues: 255-263, 'E', 265-281 <MIN1>  
 A;Note: this variant is designated albumin Herborn  
 R;Minichioti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
 Biochim. Biophys. Acta 1119, 233-238, 1992  
 A;Title: Two albumins with identical electrophoretic mobility are produced by diffe  
 A;Reference number: S21078; MUID:92190239; PMID:1347703  
 A;Accession: S21078  
 A;Molecule type: protein  
 A;Residues: 354-356, 'K', 358-378 <MIN2>  
 A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported.  
 R;He, X.M.; Carter, D.C.  
 Nature 359, 209-215, 1992  
 A;Title: Atomic structure and chemistry of human serum albumin.  
 A;Reference number: A46756; MUID:92334427; PMID:1630489  
 A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R;Brown, J.R.; Shockley, P.; Behrens, P.Q.  
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40  
 A;Reference number: A94442  
 A;Contents: annotation; three-dimensional structure and disulfide bonds  
 R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A;Title: Disulfide bonds in human serum albumin.  
 A;Reference number: A90930  
 A;Contents: annotation; disulfide bonds  
 R;Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A;Reference number: A90299; MUID:78186830; PMID:656055  
 A;Contents: annotation; bilirubin-binding site  
 R;Peters, T.; Reed, R.G.  
 In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoelholm, I., eds., 11-20,  
 A;Title: Serum albumin: conformation and active sites.  
 A;Reference number: A94408  
 A;Contents: annotation; binding sites  
 R;Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene:  
 A;Reference number: A90028; MUID:83279982; PMID:6192711  
 A;Contents: annotation; gene position  
 R;Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid  
 A;Reference number: A46755; MUID:76257808; PMID:955075  
 A;Contents: annotation  
 A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)  
 R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phospho  
 A;Reference number: A56294; MUID:92183881; PMID:1544460  
 A;Contents: annotation  
 A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in l  
 atase activity  
 C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C;Comment: A large number of variants of human serum albumin have been described.  
 C;Genetics:  
 A;Gene: GDB:ALB  
 A;Cross-references: GDB:118990; OMIM:103600  
 A;Map position: 4q11-4q13  
 C;Superfamily: serum albumin; serum albumin repeat homology  
 C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-609/Product: serum albumin #status experimental <MPT>  
 F:29-202/Domain: serum albumin repeat homology <SAL>



F:166-174/Product: kinetensin #status experimental <KIP>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 71; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14  
|||||:|||||  
Db 486 VLHEKTPVSDRVTK 499

## RESULT 2

serum albumin precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: J04660; S57632  
R:Hilger, C.; Grigioni, F.; Kentges, F.  
Gene 169, 295-296, 1996

A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A:Reference number: J04660; MUID:96194824; PMID:8647469

A:Accession: J04660  
A:Molecule type: mRNA

A:Residues: 1-608 <H12>  
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485

A:Experimental source: liver  
A:Comment: This protein is the major protein component in plasma. It functions as a multi-

chain has 35 conserved cysteine residues.  
C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: liver; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRP>  
F:25-608/Product: serum albumin #status predicted <MAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 94.4%; Score 67; DB 2; Length 608;  
Best Local Similarity 92.9%; Pred. No. 0.0042;  
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14  
|||||:|||||  
Db 486 VLHEKTPVSDRVTK 499

## RESULT 3

serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.

Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin

A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote

A:Accession: A47391  
A>Status: preliminary

A:Molecule type: mRNA; protein  
A:Residues: 1-600 <MAT>

A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295

A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:128281)

C:Superfamily: serum albumin; serum albumin repeat homology  
F:21-194/Domain: serum albumin repeat homology <SA1>

F:213-386/Domain: serum albumin repeat homology <SA2>  
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 90.1%; Score 64; DB 2; Length 600;  
Best Local Similarity 85.7%; Pred. No. 0.0015;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14  
|||||:|||||  
Db 478 VLHEKTPVSEKVTK 491

## RESULT 4

## ABPGS

serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S01382; A61006

R>Weinstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 9045, 1988

A:Title: Nucleotide sequence of porcine liver albumin.

A:Reference number: S01382; MUID:89016582; PMID:3174440

A:Accession: S01382

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-605 <WEI>

A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798

R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.

J. Bone Miner. Res. 4, 235-241, 1989

A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of minera

A:Reference number: A61006; MUID:89269769; PMID:2728927

A:Accession: A61006

A:Molecule type: protein

A:Residues: 23-51, 'X', '53-54', 'XXGY', '146', 'E', '148', 'E', '150-151', 'XVN', '155' <LIM>

A:Experimental source: dental enamel

A:Note: albumin and other serum proteins are also found in bone

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membr;

C:Superfamily: serum albumin; serum albumin repeat homology

F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>

F:17-22/Domain: propeptide #status predicted <PRO>

F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>

F:218-391/Domain: serum albumin repeat homology <SA2>

F:410-589/Domain: serum albumin repeat homology <SA3>

F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,

F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 90.1%; Score 64; DB 1; Length 605;  
Best Local Similarity 85.7%; Pred. No. 0.0015;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14  
|||||:|||||  
Db 483 VLHEKTPVSEKVTK 496

## RESULT 5

## ABPOS

serum albumin precursor [validated] - bovine

N:Alternate names: 67K protein; preproalbumin

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000

C:Accession: A38885; A36401; K91258; B60808; S10780; D45800; A26693; A90309; A91458; A9

R:Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.

submitted to the EMBL Data Library, August 1991

A:Description: Bovine serum albumin: cDNA sequence and expression.

A:Reference number: A38885

A:Accession: A38885

A:Molecule type: mRNA

A:Residues: 1-607 <HOL>

A:Cross-references: EMBL:M72215

R:Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.

Biochem. Biophys. Res. Commun. 173, 639-646, 1990

C;Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
A;Reference number: A36401; MUID:91083649; PMID:260975  
A;Accession: A36401  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-189, 'B', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HR>  
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979  
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A;Reference number: A91258; MUID:80024278; PMID:488109  
A;Accession: A91258  
A;Molecule type: protein  
A;Residues: 1-32 <MAG>  
R;Haeh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A;Title: Electrophoretic filter from an analytical isoelectrofocusing gel  
A;Reference number: A60808; MUID:88267456; PMID:3389500  
A;Accession: B60808  
A;Molecule type: protein  
A;Residues: 25-41 <HSI>  
R;Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin  
A;Reference number: S10780; MUID:90336641; PMID:2379503  
A;Accession: S10780  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>  
R;Caraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: D45800  
A;Molecule type: protein  
A;Residues: 163-172 <CAR>  
R;Caraway, R.E.; Mitra, S.P.; Cochrane, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A;Title: Structure of a biologically active neurotensin-related peptide obtained from peptidase  
A;Reference number: A26693; MUID:87194805; PMID:2437111  
A;Accession: A26693  
A;Molecule type: protein  
A;Residues: 163-172, 'L', 'CA2'  
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A;Title: Sequence of residues 400-403 of bovine serum albumin.  
A;Reference number: A90309; MUID:82023364; PMID:7283978  
A;Accession: A90309  
A;Molecule type: protein  
A;Residues: 402-433 <REE>  
R;Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A;Title: Structure of bovine serum albumin.  
A;Reference number: A91458  
A;Accession: A91458  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'B', 191-194, 'A', 196-213, 'T', 215-288, 'E'  
R;Brown, J.R.  
submitted to the Atlas, April 1975  
A;Reference number: A94551  
A;Accession: A94551  
A;Molecule type: protein  
A;Residues: 190-195 <BR2>  
R;Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A;Reference number: A91457  
A;Contents: annotation; disulfide bonds  
R;Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC 3.4.21.3)  
A;Reference number: S55232; MUID:95031935; PMID:7945219  
A;Accession: S55232  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 529-536; 569-572 <WER>  
C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; copper binding; duplication; plasma  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-607/Product: serum albumin #status experimental <MPT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, .

Query Match 90.1%; Score 64; DB 1; Length 607;  
Best Local Similarity 85.7%; Pred. NO. 0.0015;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 485 VLHEKTPVSEKVK 498  
|||||:|:|

RESULT 6  
ABSHS  
serum albumin precursor - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: S06936  
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A;Reference number: S06936; MUID:90098888; PMID:2602160  
A;Accession: S06936  
A;Molecule type: mRNA  
A;Residues: 1-607 <BR>  
A;Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, steroid hormones (weak bonds with these hormones promote their transfer across the membrane) and hormones (weak bonds with these hormones promote their transfer across the membrane).  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRO>  
F;25-607/Product: serum albumin #status predicted <MAT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, .  
F;263/Binding site: bilirubin (Lys) #status predicted

Query Match 90.1%; Score 64; DB 1; Length 607;  
Best Local Similarity 85.7%; Pred. NO. 0.0015;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 485 VLHEKTPVSEKVK 498  
|||||:|:|

RESULT 7  
ABRTS  
serum albumin precursor - rat  
N;Alternate names: preproalbumin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-May-1979 #sequence\_revision 31-May-1979 #text\_change 22-Jun-1999  
C;Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233  
R;Sargent, T.D.; Yang, M.; Bonner, J.  
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981  
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
A;Reference number: A93872; MUID:81223722; PMID:7017712  
A;Accession: A93872  
A;Molecule type: mRNA  
A;Residues: 1-608 <SAR>  
A;Cross-references: GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628  
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.  
J. Biol. Chem. 252, 6846-6855, 1977

A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis  
A;Reference number: A92211; MUID:77249657; PMID:893447  
A;Note: Cleavages during protein maturation  
A;Accession: A92211  
A;Molecule type: protein  
A;Residues: 1-38 <STR>  
R;Isemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage  
A;Reference number: A91946; MUID:78109429; PMID:564345  
A;Accession: A91946  
A;Molecule type: protein  
A;Residues: 25-222 <ISI>  
R;Isemura, S.; Ikenaka, T.  
J. Biochem. 79, 1183-1196, 1976  
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid  
A;Reference number: A91940; MUID:76260153; PMID:956149  
A;Accession: A91940  
A;Molecule type: protein  
A;Residues: 223-288;572-608 <IS2>  
A;Note: 262-Leu was also found  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 36, 3483-3486, 1976  
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A;Reference number: A90758; MUID:79001617; PMID:80265  
A;Contents: annotation; copper binding  
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: A45800  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 166-173 <CAR>  
R;Heard, J.  
Mol. Cell. Biol. 7, 2425-2434, 1987  
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved  
A;Reference number: I57621; MUID:87286876; PMID:3475566  
A;Accession: I57621  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-5 <RES>  
A;Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412  
C;Superfamily: serum albumin; duplication; metal binding; plasma  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-608/Product: serum albumin #status experimental <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status experimental  
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

Query Match 90.1%; Score 64; DB 1; Length 608;  
Best Local Similarity 85.7%; Pred. No. 0.0015;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 486 VLHEKTPVSEKVTK 499  
|||||:|||||

RESULT 8  
ABOS  
serum albumin precursor - horse  
C;Species: Equus caballus (domestic horse)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
A;Accession: S34053  
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigs, P.D.; Carter, D.C.  
Eur. J. Biochem. 215, 205-212, 1993  
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A;Reference number: S34053; MUID:93345495; PMID:8344282

A;Accession: S34053  
A;Molecule type: mRNA  
A;Residues: 1-607 <HOA>  
A;Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membr.  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRO>  
F;25-607/Product: serum albumin #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,  
F;263/Binding site: bilirubin (Lys) #status predicted

Query Match 88.7%; Score 63; DB 1; Length 607;  
Best Local Similarity 78.6%; Pred. No. 0.0023;  
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 485 VLHEKTPVSEKVTK 498  
|||||:|||||

RESULT 9  
JCS838  
albumin - Mongolian jird  
C;Species: Meriones unguiculatus (Mongolian jird)  
C;Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
A;Accession: JCS838  
R;Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.  
DNA Res. 4, 351-354, 1997  
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in th  
A;Reference number: JCS838; MUID:98116663; PMID:9455485  
A;Accession: JCS838  
A;Molecule type: mRNA  
A;Residues: 1-609 <YOS>  
A;Cross-references: DBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278  
A;Experimental source: liver  
C;Superfamily: serum albumin; serum albumin repeat homology  
F;222-395/Domain: serum albumin repeat homology <SA2>

Query Match 84.5%; Score 60; DB 2; Length 609;  
Best Local Similarity 78.6%; Pred. No. 0.008;  
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 487 LLHEKTPVSEQVTK 500  
|||||:|||||

RESULT 10  
A05139  
serum albumin - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 05-Jun-1987 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000  
A;Accession: A05139; I48638  
R;Minghetti, P.P.; Law, S.W.; Dugaiczky, A.  
Mol. Biol. Evol. 2, 347-358, 1985  
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudo;  
A;Reference number: A93055; MUID:88216123; PMID:2452956  
A;Accession: A05139  
A;Molecule type: mRNA  
A;Residues: 1-418 <MIN>  
A;Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765  
R;Boccaccio, C.; Deschattrette, J.; Meunier-Rotival, M.  
Gene 88, 181-186, 1990  
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in th  
A;Reference number: I48638; MUID:90269606; PMID:1971802  
A;Accession: I48638

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 379-453 <BOC>

A;Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334

C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; duplication; metal binding; Plasma

F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>

F;123-296/Domain: serum albumin repeat homology <SA2>

F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 83.1%; Score 59; DB 2; Length 453;

Best Local Similarity 78.6%; Pred. No. 0.0088;

Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14

Db 388 LLHEKTPVSDRVTK 401

RESULT 11

D83640

Hypothetical protein PA0049 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C;Accession: D83640

R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:2043737; PMID:10984043

A;Accession: D83640

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-553 <STO>

A;Cross-references: GB:AE004444; GB:AE004091; NID:G9945863; PIDN:AG034339.1; GSPDB:GN001

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA0049

Query Match 60.6%; Score 43; DB 2; Length 553;

Best Local Similarity 42.9%; Pred. No. 9.2;

Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14

Db 71 VTHDQTPLEIRVK 84

RESULT 12

E71937

flagellar biosynthesis protein - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Sep-1999

C;Accession: E71937

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Malls, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: E71937

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-733 <ARN>

A;Cross-references: GB:AE001473; GB:AE001439; NID:G4154910; PIDN:AAD05964.1; PID:G415491

A;Experimental source: strain J99

C;Genetics:

A;Gene: flaA

C;Superfamily: regulatory protein lcrD

Query Match 60.6%; Score 43; DB 2; Length 733;

Best Local Similarity 53.8%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 13

Db 577 LLHEKIPKIDMLT 589

RESULT 13

A64650

flagellar biosynthesis protein - Helicobacter pylori (strain 26695)

C;Species: Helicobacter pylori

C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 24-Sep-1999

C;Accession: A64650

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L

Nature 388, 533-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C

A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: A64650

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-733 <ROM>

A;Cross-references: GB:AE000612; GB:AE000511; NID:G2314188; PIDN:AAD08087.1; PID:G231411

C;Superfamily: regulatory protein lcrD

Query Match 60.6%; Score 43; DB 2; Length 733;

Best Local Similarity 53.8%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 13

Db 577 LLHEKIPKIDMLT 589

RESULT 14

S63404

hypothetical protein YNR071c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein N3605

C;Species: Saccharomyces cerevisiae

C;Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 19-Apr-2002

C;Accession: S63404

R;Andre, B.; Iraqi Housaini, I.; Urrestarazu, L.A.; Viszers, S.

submitted to the Protein Sequence Database, April 1996

A;Reference number: S62920

A;Accession: S63404

A;Molecule type: DNA

A;Residues: 1-342 <AND>

A;Cross-references: EMBL:Z71686; NID:G1302605; PID:G1302606; GSPDB:GN00014; MIPS:YNR071c

A;Experimental source: strain S288C

C;Genetics:

A;Gene: MIPS:YNR071c

A;Cross-references: SGD:S0005354

A;Map position: 14R

C;Superfamily: aldose 1-epimerase

Query Match 56.3%; Score 40; DB 2; Length 342;

Best Local Similarity 80.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSD 10

Db 231 VLHEDTPVFD 240

RESULT 15

B90289

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C;Accession: B90289

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: B90289  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-604 <KUR>  
A;Cross-references: GB:AE006641; NID:gl3814539; PIDN:AAK41569.1; GSPDB:GN00155  
C;Genetics:  
A;Gene: SSO1331

Query Match 56.3%; Score 40; DB 2; Length 604;  
Best Local Similarity 61.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13  
:|||||  
Db 6 ILHETTQVKDLVT 18

Search completed: April 19, 2004, 12:02:32  
Job time : 2.70637 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 1.02124 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_462\_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 71    | 100.0       | 609    | 1     | ALBU_HUMAN         |
| 2          | 67    | 94.4        | 608    | 1     | P49822 canis famil |
| 3          | 67    | 94.4        | 608    | 1     | P49044 felis silve |
| 4          | 64    | 90.1        | 600    | 1     | Q28522 macaca mula |
| 5          | 64    | 90.1        | 605    | 1     | P08835 sus scrofa  |
| 6          | 64    | 90.1        | 607    | 1     | P02769 bos taurus  |
| 7          | 64    | 90.1        | 607    | 1     | P14639 ovis aries  |
| 8          | 64    | 90.1        | 608    | 1     | P49065 oryctolagus |
| 9          | 64    | 90.1        | 608    | 1     | P02770 rattus norv |
| 10         | 63    | 88.7        | 607    | 1     | P35747 equus cabal |
| 11         | 60    | 84.5        | 609    | 1     | Q35090 meriones un |
| 12         | 59    | 83.1        | 608    | 1     | P07724 mus musculu |
| 13         | 43    | 60.6        | 733    | 1     | Q92m40 helicobacte |
| 14         | 43    | 60.6        | 733    | 1     | O06758 helicobacte |
| 15         | 42    | 59.2        | 905    | 1     | Q872q4 pseudomonas |
| 16         | 40    | 56.3        | 342    | 1     | P53757 saccharomyc |
| 17         | 39    | 54.9        | 346    | 1     | Q5p7y5 xestha c-ni |
| 18         | 39    | 54.9        | 728    | 1     | P44823 halobacteri |
| 19         | 38    | 53.5        | 169    | 1     | P27355 methylosinu |
| 20         | 38    | 53.5        | 539    | 1     | FUT2 ARATH         |
| 21         | 37    | 52.1        | 413    | 1     | HEMO HYACE         |
| 22         | 37    | 52.1        | 438    | 1     | SPH1 STAHY         |
| 23         | 37    | 52.1        | 505    | 1     | YNN2 CAEEL         |
| 24         | 37    | 52.1        | 661    | 1     | YBBJ SCHEO         |
| 25         | 37    | 52.1        | 900    | 1     | IF38 ARATH         |
| 26         | 37    | 52.1        | 935    | 1     | IF38 MEDTR         |
| 27         | 37    | 52.1        | 1353   | 1     | PUR2 DROME         |
| 28         | 36    | 50.7        | 256    | 1     | Y9G8 BRAJA         |
| 29         | 36    | 50.7        | 275    | 1     | LMEN-STRLN         |
| 30         | 36    | 50.7        | 331    | 1     | DP3Z MYOGE         |
| 31         | 36    | 50.7        | 514    | 1     | GPNI-SHEON         |
| 32         | 36    | 50.7        | 601    | 1     | ZF37 RAT           |
| 33         | 36    | 50.7        | 643    | 1     | GYRB_MYCCAS        |

#### ALIGNMENTS

##### RESULT 1

ALBU\_HUMAN STANDARD; PRT; 609 AA.  
 ID AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCB:TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86196112; PubMed=3009475;  
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
 RA Beattie W.G., Dugaiczky A.;  
 RT "Molecular structure of the human albumin gene is revealed by  
 RT nucleotide sequence within q11-22 of chromosome 4.";  
 RL J. Biol. Chem. 261:6747-6757(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
 RX MEDLINE=82081882; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.;  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli.";  
 RL Nucleic Acids Res. 9:6103-6114 (1981).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
 RX MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.;  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75 (1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
 RC TISSUE=fetal liver;  
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
 RT "Functional prediction of the coding sequences of 121 new genes  
 RT deduced by analysis of cDNA clones from human fetal liver.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
 RA Huang M.C., Wu H.T.;  
 RT "The cDNA sequences of human serum albumin.";  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.

34 36 50.7 734 1 PSAB CVACA Q9TLQ6 cyanidium c  
 35 36 50.7 824 1 DD20\_HUMAN O3UH16 homo sapien  
 36 36 50.7 964 1 DPOL\_CBBPV P30319 choristoneu  
 37 35 49.3 154 1 PHBS\_STAEP Q8CNS8 staphylococ  
 38 35 49.3 183 1 NUPM\_NEUCR P21976 neurospora  
 39 35 49.3 255 1 AQOD\_BACSU P35146 bacillus su  
 40 35 49.3 271 1 AQP2\_MOUSE P56402 mus musculu  
 41 35 49.3 271 1 AQP2\_RAT P34080 rattus norv  
 42 35 49.3 273 1 MADE\_STAAM Q99AX5 staphylococ  
 43 35 49.3 321 1 YBAX\_ECOLI P76254 escherichia  
 44 35 49.3 327 1 SYFA\_PASMU P57860 pasteurella  
 45 35 49.3 328 1 SDHL\_HUMAN P20132 homo sapien

RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Teisgold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RX SEQUENCE OF 25-609.  
RP MEDLINE=76187907; PubMed=1225573;  
RA Meloun B., Moravek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RX SEQUENCE OF 25-609.  
RP Brown J.R., Shockley P., Behrens P.Q.;  
RA (In) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RX SEQUENCE OF 1-455 FROM N.A.  
RP TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RX SEQUENCE OF 1-26 FROM N.A.  
RP MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RX SEQUENCE OF 222-229.  
RP MEDLINE=76257808; PubMed=955075;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylleucic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RX SEQUENCE OF 25-44 AND 480-499.  
RP TISSUE=Heart;  
RC MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RX DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RX BILIRUBIN-BINDING SITES.  
RA MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).

RN [16]  
RX VARIANT CANTERBURY ASN-337.  
RA MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RX VARIANTS NAG-2 AND NAG-3.  
RA MEDLINE=88068523; PubMed=3479777;  
RA Takehashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RX VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RA MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese alloalbumins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RX VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RA MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RX DESCRIPTION OF VARIANT REDHILL.  
RA MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
RT human serum albumin whose precursor has an aberrant signal peptidase  
RT cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RX VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RA MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
RT Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RX VARIANT VENEZIA.  
RA MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
RT carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RX VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RA MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
RN [24]  
RX VARIANT CASEROOK ASN-518.  
RA MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
RT albumin: albumin Casbrook (494 Asp-->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RX VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RA MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two alloalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 71; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.9e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 486 VLHEKTPVSDRVTK 499

RESULT 2  
ALBU CANFA STANDARD; PRT; 608 AA.  
AC P49822; Q77705, Q9TS24;  
DT 01-DEC-1996 (Rel. 34, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Can f 3).  
GN ALB.

OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Beagle; TISSUE=Liver;  
RA Hilger C.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Pandjaitan B., Swoboda I., Brandejesky-Pichler F., Rumpold H.,  
RT "Escherichia coli expression and purification of recombinant dog  
albumin, a cross-reactive animal allergen."  
RL J. Allergy Clin. Immunol. 105:279-285(2000).  
RN [3]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=75011422; PubMed=4414013;  
RA Dixon J.W., Sarkar B.;  
RT "Isolation, amino acid sequence and copper(II)-binding properties of  
peptide (1-24) of dog serum albumin."  
RL J. Biol. Chem. 249:5872-5877(1974).  
RN [4]  
RP SEQUENCE OF 25-38.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
dog heart proteins."  
RL Electrophoresis 18:2795-2802(1997).  
RN [5]  
RP SEQUENCE OF 215-478 FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=94201492; PubMed=7512102;  
RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,  
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
RT "Molecular characterization of dog albumin as a cross-reactive  
allergen."  
RL J. Allergy Clin. Immunol. 93:614-627(1994).  
RN [6]  
RP -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
RN [7]  
RP -1- SUBCELLULAR LOCATION: Secreted.  
RN [8]  
RP -1- TISSUE SPECIFICITY: Plasma.  
RN [9]  
RP -1- ALLERGEN: Causes an allergic reaction in human.  
RN [10]  
RP -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
RN [11]  
RP -1- SIMILARITY: Contains 3 albumin domains.  
RN [12]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration -  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

CC -----  
CC EMBL; AJ133489; CAB64867.1; -;  
CC EMBL; Y17737; CAA76841.1; -;  
CC EMBL; S72946; AAB30434.1; -;  
CC HSP; P02769; I57B;  
CC HSC-2DPAGE; P49822; DOG.  
CC InterPro; IPR000264; Serum\_albumin.  
CC Pfam; PF00273; transport\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin; 1.  
CC SMART; SMC0103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
CC KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
CC FT SIGNAL 1 18  
CC FT PROPEP 19 24  
CC FT CHAIN 25 608  
CC FT DOMAIN 25 205  
CC FT DOMAIN 212 397  
CC FT DOMAIN 404 595  
CC FT METAL 27 27  
CC FT DISULFID 77 86  
CC FT DISULFID 99 115  
CC FT DISULFID 114 125  
CC FT DISULFID 148 193  
CC FT DISULFID 192 201  
CC FT DISULFID 224 270  
CC FT DISULFID 269 277  
CC FT DISULFID 289 303  
CC FT DISULFID 302 313  
CC FT DISULFID 340 385  
CC FT DISULFID 384 393  
CC FT DISULFID 416 462  
CC FT DISULFID 461 472  
CC FT DISULFID 485 501  
CC FT DISULFID 500 511  
CC FT DISULFID 538 583  
CC FT DISULFID 582 591  
CC FT CONFLICT 1 26  
CC FT CONFLICT 146 146  
CC FT CONFLICT 206 206  
CC FT CONFLICT 349 349  
CC FT CONFLICT 359 359  
CC FT CONFLICT 448 448  
CC FT CONFLICT 474 474  
CC FT CONFLICT 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;  
CC SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;  
CC REF. 2)  
CC A -> R (IN REF. 2).  
CC I -> T (IN REF. 2).  
CC V -> A (IN REF. 2).  
CC S -> A (IN REF. 1).  
CC V -> VV (IN REF. 5).  
CC D -> E (IN REF. 1).  
CC MKVTFISLFLFSSAYSRGLVREA -> MDT (IN  
CC REF. 2).  
CC Query Match 94.4%; Score 67; DB 1; Length 608;  
CC Best Local Similarity 92.9%; Pred. No. 0.00035;  
CC Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 486 VLHEKTPVSDRVTK 499

RESULT 3  
ALBU FELCA STANDARD; PRT; 608 AA.  
ID ALBU FELCA  
AC P49064;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Fel d 2).  
GN ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.



OX NCBI\_TaxID=9685;  
RN SEQUENCE FROM N.A.  
RP MEDLINE=96194824; PubMed=8647469;  
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;  
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";  
RL Gene 169:295-296(1996).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X84842; CAA59279.1; -  
DR PIR; JC4660; S57632.  
DR HSSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR PRODom; PD002486; Serum albumin; 1.  
DR SMART; SMO0103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
FT SIGNAL 1 18  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 608 SERUM ALBUMIN.  
FT DOMAIN 25 205 ALBUMIN 1.  
FT DOMAIN 212 397 ALBUMIN 2.  
FT DOMAIN 404 595 ALBUMIN 3.  
FT METAL 27 27 COPPER.  
FT DISULFID 77 86 BY SIMILARITY.  
FT DISULFID 99 115 BY SIMILARITY.  
FT DISULFID 114 125 BY SIMILARITY.  
FT DISULFID 148 193 BY SIMILARITY.  
FT DISULFID 132 201 BY SIMILARITY.  
FT DISULFID 224 270 BY SIMILARITY.  
FT DISULFID 269 277 BY SIMILARITY.  
FT DISULFID 289 303 BY SIMILARITY.  
FT DISULFID 302 313 BY SIMILARITY.  
FT DISULFID 340 395 BY SIMILARITY.  
FT DISULFID 384 393 BY SIMILARITY.  
FT DISULFID 416 462 BY SIMILARITY.  
FT DISULFID 461 472 BY SIMILARITY.  
FT DISULFID 485 501 BY SIMILARITY.  
FT DISULFID 500 511 BY SIMILARITY.  
FT DISULFID 538 583 BY SIMILARITY.  
FT DISULFID 582 591 BY SIMILARITY.  
FT SEQUENCE 608 AA; 68659 MW; 07B629CAC3F60E5F CRC64;  
SQ  
Query Match 94.4%; Score 67; DB 1; Length 608;  
Best Local Similarity 92.9%; Pred. No. 0.00035;  
Matches 13; Conservative 1; Mismatches 0; Gaps 0;  
OY 1 VLHEKTPVSDRVTK 14  
||| ||||| |||||  
Db 486 VLHEKTPVSDRVTK 499  
RESULT 4  
ALBU\_MACMU  
ID \_ALBU\_MACMU STANDARD; PRT; 600 AA.  
AC Q28522;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor (Fragment).  
GN AUB.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93211971; PubMed=8460152;  
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,  
RA Dwulet J., Putnam F.W.;  
RT "cDNA and protein sequence of polymorphic macaque albumins that differ  
RT in bilirubin binding";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M90463; AAA36906.1; -  
DR PIR; A47391; A47391.  
DR HSSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR PRODom; PD002486; Serum albumin; 1.  
DR SMART; SMO0103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 10 BY SIMILARITY.  
FT PROPEP 11 16 BY SIMILARITY.  
FT CHAIN 17 600 SERUM ALBUMIN.  
FT DOMAIN 17 197 ALBUMIN 1.  
FT DOMAIN 204 389 ALBUMIN 2.  
FT DOMAIN 396 587 ALBUMIN 3.  
FT METAL 19 19 COPPER (BY SIMILARITY).  
FT BINDING 256 256 BILIRUBIN (POTENTIAL).  
FT DISULFID 69 78 BY SIMILARITY.  
FT DISULFID 91 107 BY SIMILARITY.  
FT DISULFID 106 117 BY SIMILARITY.  
FT DISULFID 140 185 BY SIMILARITY.  
FT DISULFID 184 193 BY SIMILARITY.  
FT DISULFID 216 262 BY SIMILARITY.  
FT DISULFID 261 269 BY SIMILARITY.  
FT DISULFID 281 295 BY SIMILARITY.  
FT DISULFID 332 377 BY SIMILARITY.  
FT DISULFID 376 385 BY SIMILARITY.  
FT DISULFID 408 454 BY SIMILARITY.  
FT DISULFID 453 464 BY SIMILARITY.  
FT DISULFID 477 493 BY SIMILARITY.  
FT DISULFID 492 503 BY SIMILARITY.  
FT DISULFID 530 575 BY SIMILARITY.  
FT DISULFID 574 583 BY SIMILARITY.  
FT SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;  
SQ

```
Query Match          90.1%; Score 64; DB 1; Length 600;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
        |||||:|||||
DB      478 VLHEKTPVSEKVK 491

RESULT 5
ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstein J.;
RT "Nucleotide sequence of porcine liver albumin.";
RN Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X12422; CAA30970.1; -
DR EMBL; M36787; AAA30988.1; -
DR PIR; S01382; ABPGS.
DR HSP; P02768; 1E7H.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; P00273; transport_prot; 3.
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; P002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1
FT PROPEP 41 16 BY SIMILARITY.
FT CHAIN 17 22 BY SIMILARITY.
FT DOMAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT METAL 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.

Query Match          90.1%; Score 64; DB 1; Length 605;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
        |||||:|||||
DB      483 VLHEKTPVSEKVK 496

RESULT 6
ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RL "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
RN [5]
RP SEQUENCE OF 1-32. PubMed=488109;
RX MEDLINE=80024278;
RA McGillivray R.T.A., Chung D.W., Davie B.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RT terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RL "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.
```

RX MEDLINE=82023364; PubMed=7283978;  
RA Read R.G., Putnam F.W., Peters T. Jr.;  
RL "Sequence of residues 400-403 of bovine serum albumin.";  
RN Biochem. J. 191:867-868(1980).  
RP SEQUENCE OF 19-28.  
RX MEDLINE=77134075; PubMed=843354;  
RA Patterson J.E., Geller D.M.;  
RL "Bovine microsomal albumin: amino terminal sequence of bovine  
RL prealbumin.";  
RN Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
RX MEDLINE=91083649; PubMed=2260975;  
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;  
RL "Rapid confirmation and revision of the primary structure of bovine  
RL serum albumin by EMS and Frit-FAB LC/MS.";  
RN Biochem. Biophys. Res. Commun. 173:639-646(1990).  
RP SEQUENCE OF 25-41.  
RX MEDLINE=98267456; PubMed=3389500;  
RA Hsieh J.C., Lin F.P., Tam M.F.;  
RL "Electroblotting onto glass-fiber filter from an analytical  
RL isoelectrofocusing gel: a preparative method for isolating proteins  
RL for N-terminal microsequencing.";  
RN Anal. Biochem. 170:1-8(1988).  
RP SEQUENCE OF 437-451.  
RA Vilbois F.;  
RL Submitted (AUG-1998) to Swiss-Prot.  
RN DISULFIDE BONDS.  
RP Structure of serum albumin: disulfide bridges.";  
RA Brown J.R.;  
RL Fed. Proc. 33:1389-1399(1974).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -1- TISSUE SPECIFICITY: Plasma.  
CC -1- ALLERGEN: Causes an allergic reaction in human.  
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -1- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; M73993; AAA51411.1; --  
DR EMBL; X59389; CAA41735.1; --  
DR EMBL; Y17769; CAA76847.1; --  
DR EMBL; AF542068; AAN1824.1; --  
DR HSSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Transport prot.; 3.  
DR PRINTS; P00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN7.3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
KW Polymorphism.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607 SERUM ALBUMIN.  
FT DOMAIN 25 204 ALBUMIN 1.  
FT DOMAIN 211 396 ALBUMIN 2.  
FT DOMAIN 403 594 ALBUMIN 3.  
FT METAL 27 27 COPPER (BY SIMILARITY).

FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 132  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
FT VARIANT 214 214  
FT CONFLICT 302 302  
FT CONFLICT 304 305  
FT CONFLICT 324 324  
FT CONFLICT 394 395  
FT CONFLICT 437 437  
FT CONFLICT 493 494  
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768595D4 CRC64;  
Query Match 90.1%; Score 64; DB 1; Length 607;  
Best Local Similarity 85.7%; Pred. No. 0.0012; Mismatches 0; Indels 0; Gaps 0;  
Matches 12; Conservative 2;  
QY 1 VLHEKTPVSDRVTK 14  
DB 485 VLHEKTPVSEKVTK 498  
|||||:|:|  
RESULT 7  
ALBU\_SHEEP STANDARD; PRT; 607 AA.  
ID ALBU\_SHEEP  
AC P14639;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID:9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=90098888; PubMed=2602160;  
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;  
RL "Nucleotide and deduced amino acid sequence of sheep serum albumin.";  
RL Nucleic Acids Res. 17:10495-10495(1989).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Plasma.  
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -1- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

```

DR EMBL; X17055; CAA34903.1; -.
DR PIR; S06936; ABSHS.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT CHAIN 19 24
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 90.1%; Score 64; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 485 VLHEKTPVSEKVTK 498

RESULT 8
ALBU RABIT
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

---

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18344; AAB58347.1; -.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT CHAIN 19 24
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 90.1%; Score 64; DB 1; Length 608;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 486 VLHEKTPVSEKVTK 499

RESULT 9
ALBU RAT
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=81233722; PubMed=7017712;
RA Sargent T.D., Yang M., Borner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77243657; PubMed=893447;

```

RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-  
RT piece. Analysis of the direct translation product of albumin  
RL messenger RNA."; J. Biol. Chem. 252:6846-6855 (1977).  
RN [3]  
RP SEQUENCE OF 25-222.  
RX Iseura S., Ikenaka T.; PubMed=564345;  
RA MEDLINE=78109429; PubMed=564345;  
RT "Amino acid sequences of fragments I and II obtained by cyanogen  
RT bromide cleavage of rat serum albumin."; J. Biochem. 83:35-48 (1978).  
RL [4]  
RN SEQUENCE OF 223-288 AND 572-608.  
RX Iseura S., Ikenaka T.; PubMed=956149;  
RA MEDLINE=76260153; PubMed=956149;  
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
RT the amino acid sequences of four fragments."; J. Biochem. 79:1183-1196 (1976).  
RL [5]  
RN SEQUENCE OF 166-174.  
RX TISSUE=Plasma;  
RA MEDLINE=87194805; PubMed=2437111;  
RT Caraway R.E., Mitra S.P., Cochran D.E.;  
RT "Structure of a biologically active neurotensin-related peptide  
RT obtained from pepsin-treated albumin(s)"; J. Biol. Chem. 262:5968-5973 (1987).  
RL [6]  
RN COPPER-BINDING.  
RX MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein."; Cancer Res. 38:3483-3486 (1978).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and  
CC blood flow (potential).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; V01222; CAA24532.1; -  
CC PIR; A93872; ABETS.  
CC HSP; P02768; 1E7B.  
CC InterPro; IPR000264; Serum\_albumin.  
CC Pfam; PF00273; transport\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608 SERUM ALBUMIN.  
FT PEPTIDE 166 174 NEUTROGENIN-RELATED PEPTIDE.  
FT DOMAIN 25 205 ALBUMIN 1.  
FT DOMAIN 212 397 ALBUMIN 2.  
FT DOMAIN 404 595 ALBUMIN 3.  
FT METAL 27 27 COPPER.  
FT DISULFID 77 86 BY SIMILARITY.  
FT DISULFID 99 115 BY SIMILARITY.  
FT DISULFID 114 125 BY SIMILARITY.

FT DISULFID 148 193 BY SIMILARITY.  
FT DISULFID 192 201 BY SIMILARITY.  
FT DISULFID 224 270 BY SIMILARITY.  
FT DISULFID 269 277 BY SIMILARITY.  
FT DISULFID 289 303 BY SIMILARITY.  
FT DISULFID 302 313 BY SIMILARITY.  
FT DISULFID 340 385 BY SIMILARITY.  
FT DISULFID 384 393 BY SIMILARITY.  
FT DISULFID 416 462 BY SIMILARITY.  
FT DISULFID 461 472 BY SIMILARITY.  
FT DISULFID 485 501 BY SIMILARITY.  
FT DISULFID 500 511 BY SIMILARITY.  
FT DISULFID 538 583 BY SIMILARITY.  
FT DISULFID 582 591 BY SIMILARITY.  
FT VARIANT 282 282  
FT CONFLICT 174 174 Y -> L (IN REF. 5).  
SQ SEQUENCE 608 AA; 68718 MW; 5B497A282411AB7 CRC64;  
Query Match 90.1%; Score 64; DB 1; Length 608;  
Best Local Similarity 85.7%; Pred. No. 0.0012;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLHEKTPVSDRVTK 14  
Db 486 VLHEKTPVSEKVTK 499  
|||||:|:|:  
RESULT 10  
ALBU\_HORSE STANDARD; PRT; 607 AA.  
ID ALBU\_HORSE  
AC P35747;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Equ c 3).  
GN ALB.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
RC TISSUE=Liver;  
RX MEDLINE=93345495; PubMed=8344282;  
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
RT at 0.27-nm resolution."; Eur. J. Biochem. 215:205-212 (1993).  
RL -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X74045; CAA52194.1; -  
CC PIR; S34053; ABHOS.  
CC HSP; P02768; 1E7B.  
CC InterPro; IPR000264; Serum\_albumin.  
CC Pfam; PF00273; transport\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin; 1.

```

DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT METAL 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 88.7%; Score 63; DB 1; Length 607;
Best Local Similarity 78.6%; Pred. No. 0.0018;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
    |||||:|||||:|||||
Db 485 VLHEKTPVSEKTK 498

RESULT 11
ALBU MERUN
ID _ALBU MERUN STANDARD; PRT; 609 AA.
AC O35090;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=9811663; PubMed=9455485;
RA Yoshida K., Seto-Onshima A., Sinochawa H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RL DNA Res. 4:351-354(1997).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

---

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB006197; BAA21765.1; -.
DR PIR; JC5838; JC5838.
DR HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PDO02486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 609
FT DOMAIN 25 206
FT DOMAIN 213 398
FT DOMAIN 405 596
FT METAL 28 28
FT METAL 78 87
FT DISULFID 100 116
FT DISULFID 115 126
FT DISULFID 149 194
FT DISULFID 193 202
FT DISULFID 225 271
FT DISULFID 270 278
FT DISULFID 290 304
FT DISULFID 303 314
FT DISULFID 341 386
FT DISULFID 385 394
FT DISULFID 417 463
FT DISULFID 462 473
FT DISULFID 486 502
FT DISULFID 501 512
FT DISULFID 539 584
FT DISULFID 583 592
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F67E7FLA48 CRC64;

Query Match 84.5%; Score 60; DB 1; Length 609;
Best Local Similarity 78.6%; Pred. No. 0.0062;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
    :|||||:|||||:|||||
Db 487 LLHEKTPVSEQVTK 500

RESULT 12
ALBU MOUSE
ID _ALBU_MOUSE STANDARD; PRT; 608 AA.
AC P07724; Q61802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALB1 OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085650; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

```

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Peole G., Quackenbush J.,  
RA Schindl L.M., Staali P., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Ioyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontecki S.,  
RA Hayashizaki Y.,  
RA Nature 409:685-690(2001).  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN (3)  
RP SEQUENCE FROM N.A.  
RX STRAIN=FVB/N; TISSUE=Liver;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Lequellano N.J., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McKernan K.J., Malek J.A., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 99-516 FROM N.A.  
RX MEDLINE=88216123; PubMed=2452956;  
RA Minghetti P.P., Law S.W., Dugaiczky A.,  
RT "The rate of molecular evolution of alpha-fetoprotein approaches that  
of pseudogenes.";  
RL Mol. Biol. Evol. 2:347-358(1985).  
RN [5]  
RP SEQUENCE OF 477-551 FROM N.A.  
RX STRAIN=BALB/C;  
RA Boccaccio C., Deschatrete J., Meunier-Rotival M.,  
RT "Empty and occupied insertion site of the truncated LINE-1 repeat  
located in the mouse serum albumin-encoding gene.";  
RL Gene 88:181-186(1990).  
RN [6]  
RP SEQUENCE OF 25-44.  
RX TISSUE=Liver;  
RA MEDLINE=93162044; PubMed=1286668;  
RA Giometti C.S., Taylor J., Tollaksen S.L.,  
RT "Mouse liver protein database: a catalog of proteins detected by two-  
dimensional gel electrophoresis.";  
RL Electrophoresis 13:970-991(1992).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AJ011413; CAA09617.1; -;  
CC EMBL; AK010025; BAB26650.1; -;  
CC EMBL; BC049971; AAB49971.1; -;  
CC EMBL; M16111; AAA37190.1; -;  
CC EMBL; X13060; CAA31458.1; -;  
CC PIR; A05139; A05139.  
CC HSSP; P02768; 1E78.  
CC SWISS-2DPAGE; P07724; MOUSE.  
CC MGD; MGI:87991; Albl.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transprot\_prot; 3.  
CC PRINTS; PRO0802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN\_3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
CC SIGNAL 1 18  
CC PROPEP 19 24  
CC CHAIN 25 608  
CC DOMAIN 25 205  
CC DOMAIN 212 397  
CC DOMAIN 404 595  
CC METAL 27 27  
CC DISULFID 77 86  
CC DISULFID 99 115  
CC DISULFID 114 125  
CC DISULFID 148 193  
CC DISULFID 192 201  
CC DISULFID 224 270  
CC DISULFID 269 277  
CC DISULFID 289 303  
CC DISULFID 302 313  
CC DISULFID 340 385  
CC DISULFID 384 393  
CC DISULFID 416 462  
CC DISULFID 461 472  
CC DISULFID 485 501  
CC DISULFID 500 511  
CC DISULFID 538 583  
CC DISULFID 582 591  
CC CONFLICT 27 27  
CC CONFLICT 33 33  
CC CONFLICT 41 41  
CC SEQUENCE 608 AA; 68692 MW; 292F7C7ED3A61B4 CRC64;  
Query Match 83.1%; Score 59; DB 1; Length 608;  
Best Local Similarity 78.6%; Pred. No. 0.0093;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLHEKTPVSDRTK 14  
Db 486 LLHEKTPVSEHTK 499  
RESULT 13  
FLHA HELPJ STANDARD; PRT; 733 AA.  
ID FLHA HELPJ  
AC Q9ZM40;  
DT 16-OCT-2001 (Rel. 40, Created)  
DF 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Flagellar biosynthesis protein flha.  
GN FLHA OR FLBA OR JHP0383.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen Helicobacter pylori.";  
RL Nature 397:176-180(1999).  
CC -!- FUNCTION: INVOLVED IN THE EXPORT OF FLAGELLUM PROTEINS.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE FHPEP (FLAGELLA/HR/INVASION PROTEINS  
CC EXPORT PORE) FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE001473; AAD05964.1; -.  
DR PIR; E71937; E71937.  
DR InterPro; IPR001712; Bact FHPEP.  
DR InterPro; IPR006301; Flha-  
DR Pfam; PF00771; FHPEP; 1  
DR PRINTS; PR00949; TYPE3IMAPROT.  
DR TIGRFAMs; TIGR01398; Flha; 1.  
DR PROSITE; PS00994; FHPEP; FALSE NEG.  
KW Flagellum; Transport; Protein transport; Transmembrane;  
KW Complete proteome.  
FT TRANSMEM 28 48 POTENTIAL.  
FT TRANSMEM 49 69 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT TRANSMEM 246 266 POTENTIAL.  
FT TRANSMEM 291 311 POTENTIAL.  
FT TRANSMEM 313 333 POTENTIAL.  
SQ SEQUENCE 733 AA; 80778 MW; EDB00DF3A2F28F0E CRC64;  
Query Match 60.6%; Score 43; DB 1; Length 733;  
Best Local Similarity 53.8%; Pred. No. 7.7;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 VLHEKTPVSDRVT 13  
DB 577 LHEKIPKIDMLT 589  
RESULT 14  
ID FLHA HELP STANDARD; PRT; 733 AA.  
AC O06758; O06759;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Flagellar biosynthesis protein flha.  
DE FLHA OR FLBA OR HP1041.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=85P;  
RA Schmitz A., Josenhans C., Suerbaum S.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

[2]  
RN RP SEQUENCE FROM N.A.  
RX STRAIN=26895 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori.";  
RL Nature 388:539-547(1997).  
RN [3]  
RP SEQUENCE OF 93-733 FROM N.A.  
RC STRAIN=ATCC 49503 / 60190;  
RA Fauconier A., Lage A., Bollen A., Godtfroid E.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INVOLVED IN THE EXPORT OF FLAGELLUM PROTEINS.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE FHPEP (FLAGELLA/HR/INVASION PROTEINS  
CC EXPORT PORE) FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE000612; AAD08087.1; -.  
DR EMBL; Y13395; CAA73825.1; -.  
DR EMBL; Y08762; CAA73007.1; -.  
DR PIR; A64650; A64650.  
DR TIGR; HP1041; -.  
DR InterPro; IPR001712; Bact FHPEP.  
DR InterPro; IPR006301; Flha-  
DR Pfam; PF00771; FHPEP; 1  
DR PRINTS; PR00949; TYPE3IMAPROT.  
DR TIGRFAMs; TIGR01398; Flha; 1.  
DR PROSITE; PS00994; FHPEP; FALSE NEG.  
KW Flagellum; Transport; Protein transport; Transmembrane;  
KW Complete proteome.  
FT TRANSMEM 28 48 POTENTIAL.  
FT TRANSMEM 49 69 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT TRANSMEM 246 266 POTENTIAL.  
FT TRANSMEM 291 311 POTENTIAL.  
FT TRANSMEM 313 333 POTENTIAL.  
FT TRANSMEM 113 113 D -> I (IN REF. 1).  
FT TRANSMEM 267 267 A -> R (IN REF. 1).  
FT TRANSMEM 305 307 LLL -> YC (IN REF. 1).  
FT TRANSMEM 309 309 A -> C (IN REF. 1).  
FT TRANSMEM 324 324 T -> A (IN REF. 1).  
FT TRANSMEM 370 370 T -> A (IN REF. 1 AND 3).  
FT TRANSMEM 381 381 L -> I (IN REF. 1).  
FT TRANSMEM 407 407 I -> Y (IN REF. 1).  
FT TRANSMEM 670 670 V -> A (IN REF. 1).  
FT TRANSMEM 670 670 V -> G (IN REF. 3).  
FT TRANSMEM 710 710 I -> V (IN REF. 3).  
SQ SEQUENCE 733 AA; 80908 MW; C15B4C9EBB3333C5 CRC64;  
Query Match 60.6%; Score 43; DB 1; Length 733;  
Best Local Similarity 53.8%; Pred. No. 7.7;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 VLHEKTPVSDRVT 13  
DB 577 LHEKIPKIDMLT 589



```

Query Match          59.2%; Score 42; DB 1; Length 905;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 HNKTPVSDR 11
      |:|:|:|
Db      592 HQKTPSDR 600

Search completed: April 19, 2004, 11:52:56
Job time : 2.02124 secs

```

GenCore version 5.1.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 5.40351 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_462\_475  
Perfect score: 71  
Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriaph: \*  
17: sp\_archaea: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description         |
|------------|-------|---------------|--------|----|---------------------|
| 1          | 71    | 100.0         | 396    | 4  | Q8iuk7 homo sapien  |
| 2          | 71    | 100.0         | 417    | 4  | Q86ygo homo sapien  |
| 3          | 67    | 94.4          | 584    | 6  | Q7ysg3 helix silve  |
| 4          | 63    | 88.7          | 608    | 5  | Q95vb7 schistosoma  |
| 5          | 59    | 83.1          | 576    | 11 | Q8c7c7 mus musculus |
| 6          | 59    | 83.1          | 608    | 11 | Q8c7h3 mus musculus |
| 7          | 46    | 64.8          | 148    | 10 | Q8gsh6 oryza sativ  |
| 8          | 43    | 60.6          | 553    | 16 | Q91783 pseudomonas  |
| 9          | 43    | 60.6          | 733    | 2  | Q86g99 helicobacte  |
| 10         | 42    | 59.2          | 905    | 16 | Q87zq4 pseudomonas  |
| 11         | 41    | 57.7          | 95     | 2  | Q93eh4 helicobacte  |
| 12         | 41    | 57.7          | 231    | 16 | Q8f9d1 leptospira   |
| 13         | 41    | 57.7          | 735    | 16 | Q7viy5 helicobacte  |
| 14         | 40    | 56.3          | 358    | 2  | Q8vut7 burkholderi  |
| 15         | 40    | 56.3          | 604    | 17 | Q97vj0 sulfolobus   |
| 16         | 40    | 56.3          | 630    | 10 | Q9sf07 arabidopsis  |

|    |    |      |      |    |        |
|----|----|------|------|----|--------|
| 17 | 40 | 56.3 | 652  | 4  | Q8N990 |
| 18 | 40 | 56.3 | 699  | 5  | O01545 |
| 19 | 40 | 56.3 | 3456 | 12 | P89201 |
| 20 | 39 | 54.9 | 63   | 16 | Q989J5 |
| 21 | 39 | 54.9 | 137  | 2  | Q8KZS5 |
| 22 | 39 | 54.9 | 385  | 16 | Q92KLO |
| 23 | 39 | 54.9 | 577  | 2  | O33047 |
| 24 | 39 | 54.9 | 682  | 17 | Q9HM85 |
| 25 | 39 | 54.9 | 753  | 16 | Q9CBE1 |
| 26 | 39 | 54.9 | 1125 | 5  | Q9NFA7 |
| 27 | 38 | 53.5 | 172  | 16 | O69637 |
| 28 | 38 | 53.5 | 172  | 16 | O7TVZ5 |
| 29 | 38 | 53.5 | 309  | 2  | P72172 |
| 30 | 38 | 53.5 | 309  | 5  | Q9NJ11 |
| 31 | 38 | 53.5 | 309  | 16 | Q9HUZ7 |
| 32 | 38 | 53.5 | 333  | 16 | Q8DTB0 |
| 33 | 38 | 53.5 | 375  | 16 | Q8A079 |
| 34 | 38 | 53.5 | 444  | 10 | Q8S9T6 |
| 35 | 38 | 53.5 | 457  | 10 | O82108 |
| 36 | 38 | 53.5 | 577  | 2  | Q9AL91 |
| 37 | 38 | 53.5 | 690  | 16 | Q97H66 |
| 38 | 38 | 53.5 | 768  | 2  | Q84FP3 |
| 39 | 38 | 53.5 | 777  | 2  | Q8GM52 |
| 40 | 38 | 53.5 | 931  | 16 | Q82K00 |
| 41 | 38 | 53.5 | 1522 | 5  | Q7YSM9 |
| 42 | 38 | 53.5 | 1771 | 3  | O870V1 |
| 43 | 37 | 52.1 | 131  | 6  | Q8HXG0 |
| 44 | 37 | 52.1 | 161  | 16 | Q8FOA5 |
| 45 | 37 | 52.1 | 216  | 3  | Q9USA7 |

ALIGNMENTS

RESULT 1  
Q8iuk7 PRELIMINARY; PRT; 396 AA.  
AC Q8iuk7;  
CT 01-11-2003 (TRENBLrel. 23, Created)  
DT 01-11-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-11-2003 (TRENBLrel. 25, Last annotation update)  
DE Similar to serum albumin precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035969; AAH35969.1; -  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0008610; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 2.  
DR PRINTS; PRO0802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 71; DB 4; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14  
DB 273 VLHEKTPVSDRVTK 286

RESULT 2  
Q86ygo

```
ID Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47350 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 71; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 294 VLHEKTPVSDRVTK 307
|||||
|||||

RESULT 3
Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R.; Swoboda I.; Bohle B.; Hauswirth A.W.; Valent P.;
RA Rumpold H.; Valente R.; Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -.
FT NON_TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 94.4%; Score 67; DB 6; Length 584;
Best Local Similarity 92.9%; Pred. No. 0.0016;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 462 VLHEKTPVSDRVTK 475
|||||
|||||

RESULT 4
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A.; Asahi H.; Stadecker M.J.; LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E54 CRC64;

Query Match 88.7%; Score 63; DB 5; Length 608;
Best Local Similarity 85.7%; Pred. No. 0.0086;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 486 VLHEKTPVSDRVTK 499
|||||
|||||

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050844; BAC34360.1; -.
DR MGD; MGI:187991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;
```

```

Query Match      83.1%; Score 59; DB 11; Length 576;
Best Local Similarity 78.6%; Pred. No. 0.043; 1; Indels 0; Gaps 0;
Matches 11; Conservative 2; Mismatches 0;

QY 1 VLHEKTPVSDRVTK 14
Db 454 LLHEKTPVSEHVTK 467

RESULT 6
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALBL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; ALBL.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match      83.1%; Score 59; DB 11; Length 608;
Best Local Similarity 78.6%; Pred. No. 0.046; 1; Indels 0; Gaps 0;
Matches 11; Conservative 2; Mismatches 0;

QY 1 VLHEKTPVSDRVTK 14
Db 486 LLHEKTPVSEHVTK 499

RESULT 7
Q8GSH6 PRELIMINARY; PRT; 148 AA.
AC Q8GSH6;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE OJ1477_F01.19 protein (P0496C02.4 protein).
GN OJ1477_F01.19 OR P0496C02.4.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1477_F01.";
```

```

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0496C02.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003833; BAC15482.1; -.
DR EMBL; AP004378; BAC16046.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 2.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK_REPEAT; 2.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
SQ SEQUENCE 148 AA; 15284 MW; 1D69B6SDE09D4364 CRC64;

Query Match      64.8%; Score 46; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 2.5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 HEKTPVSDRVTK 14
Db 113 HEKTPVSDRVTK 124

RESULT 8
Q9I783 PRELIMINARY; PRT; 553 AA.
AC Q9I783;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein PA0049.
GN PA0049.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Muzoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Geltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004444; AAG03439.1; -.
DR PIR; D83640; D83640.
DR GO; GO:0004219; F:pyroglutamy-peptidase I activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000816; Peptidase_C15.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 553 AA; 60113 MW; BE9B7E723D46FB3B CRC64;

Query Match      60.6%; Score 43; DB 16; Length 553;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 71 VTHDQTPLEAERIVK 84

RESULT 9
Q9S6G9 PRELIMINARY; PRT; 733 AA.
AC Q9S6G9;
```

```

DR 01-MAY-2000 (TrEMBLrel. 13, Created)
DR 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DR 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Flagellar biosynthesis/regulation protein FLBA.
GN FLBA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43504;
RX MEDLINE=99214098; PubMed=10198012;
RA McGee D.J., May C.A., Garner R.M., Himpel J.M., Mobley H.L.T.;
RT "Isolation of Helicobacter pylori genes that modulate urease
activity."
RL J. Bacteriol. 181:2477-2484 (1999).
DR EMBL; AF125197; AAD27689.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR001712; Bact_FHIPPEP.
DR InterPro; IPR006301; FLBA-.
DR Pfam; PF00771; FHIPPEP.1.
DR PRINTS; PR00949; TYP3IAPROT.
DR TIGRFAMs; TIGR01398; FLBA.1.
SQ SEQUENCE 733 AA; 80883 MW; 6E8817AB4268129B CRC64;

Query Match 60.6%; Score 43; DB 2; Length 733;
Best Local Similarity 53.8%; Pred. No. 43;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13
:|||||:|:|:|
DB 577 LLHEKIPKDMLT 599

RESULT 10
Q87ZQ4 PRELIMINARY; PRT; 905 AA.
ID Q87ZQ4;
AC Q87ZQ4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I, G subunit.
GN NUOG OR PSPT03370.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016968; AA056848.1; -.
DR TIGR; PST03370; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR009010; Asp decarb fold.
DR InterPro; IPR000283; Complex1_75K.
DR InterPro; IPR001041; Ferridoxin.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006963; Molybdop_Fe4S4.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00384; molybdopterin; 1.
DR Pfam; PF04879; Molybdop_Fe4S4; 1.

Query Match 57.7%; Score 41; DB 2; Length 95;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13
:|||||:|:|:|
DB 51 LLHEKVPFKDMPT 63

RESULT 12
Q8F9B1 PRELIMINARY; PRT; 231 AA.
ID Q8F9B1;
AC Q8F9B1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA0284.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011216; AAN47483.1; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 231 AA; 26471 MW; 15024FDS6DE290D9 CRC64;

```

[illegible]

ng

Italy";  
Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725 (1990).  
[22]  
VARIANT VENEZIA.  
MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
RT carboxyl-terminal variants of human serum albumin.";  
Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963 (1991).  
[23]  
VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RA KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuoka Y.-I., Anaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857 (1991).  
[24]  
VARIANT CASEBOOK ASN-518.  
MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
RT albumin: albumin Casebrook (494 Asp-->Asn).";  
Biochim. Biophys. Acta 1097:49-54 (1991).  
[25]  
VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two allalbumins with identical electrophoretic mobility are produced  
Query Match 100.0%; Score 33; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HGDLL 6  
Db 271 HGDLL 276  
RESULT 15  
ALBU MERUN  
ID ALBU MERUN STANDARD; PRT; 609 AA.  
AC Q35090;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
OC Meriones.  
OX NCBI\_TaxID=10047;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=MGS IDR; TISSUE=Liver;  
RX MEDLINE=98116663; PubMed=9455485;  
RA Yoshida K., Seto-Oshima A., Sinohara H.;  
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
RNA Res. 4:351-354 (1997).  
CC -|- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- TISSUE SPECIFICITY: Plasma.  
CC -|- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -|- SIMILARITY: Contains 3 albumin domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
CC EMBL; AB006197; BAA21765.1; -;  
CC PIR; JCS838; JCS838.  
CC HSSP; P02768; 1E7B.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport p1ct; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC Prodom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 609 SERUM ALBUMIN.  
FT DOMAIN 25 206 ALBUMIN 1.  
FT DOMAIN 213 398 ALBUMIN 2.  
FT METAL 405 596 ALBUMIN 3.  
FT METAL 28 28 COPPER.  
FT DISULFID 78 87 BY SIMILARITY.  
FT DISULFID 100 116 BY SIMILARITY.  
FT DISULFID 115 126 BY SIMILARITY.  
FT DISULFID 149 194 BY SIMILARITY.  
FT DISULFID 193 202 BY SIMILARITY.  
FT DISULFID 225 271 BY SIMILARITY.  
FT DISULFID 270 278 BY SIMILARITY.  
FT DISULFID 290 304 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 341 386 BY SIMILARITY.  
FT DISULFID 385 394 BY SIMILARITY.  
FT DISULFID 417 463 BY SIMILARITY.  
FT DISULFID 462 473 BY SIMILARITY.  
FT DISULFID 486 502 BY SIMILARITY.  
FT DISULFID 501 512 BY SIMILARITY.  
FT DISULFID 539 584 BY SIMILARITY.  
FT DISULFID 583 592 BY SIMILARITY.  
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F67EFA48 CRC64;  
Query Match 100.0%; Score 33; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HGDLL 6  
Db 272 HGDLL 277  
Search completed: April 19, 2004, 11:52:50  
Job time : 0.437673 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.31579 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLLLE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 33    | 100.0       | 123    | 11 Q7TQM6 | Q7TQM6 meriones un  |
| 2          | 33    | 100.0       | 129    | 6 Q8WMJ3  | Q8wmj3 macaca mulla |
| 3          | 33    | 100.0       | 194    | 6 Q8WMH8  | Q8wmh8 macaca mulla |
| 4          | 33    | 100.0       | 211    | 6 O19044  | O19044 cryptolagus  |
| 5          | 33    | 100.0       | 250    | 16 Q91612 | Q91612 pseudomonas  |
| 6          | 33    | 100.0       | 251    | 2 Q84712  | Q84712 pseudomonas  |
| 7          | 33    | 100.0       | 283    | 2 Q8RBV8  | Q8rbv8 pseudomonas  |
| 8          | 33    | 100.0       | 379    | 4 Q8N417  | Q8n417 homo sapien  |
| 9          | 33    | 100.0       | 379    | 11 Q8C7Z5 | Q8c7z5 mus musculus |
| 10         | 33    | 100.0       | 417    | 4 Q86YGO  | Q86ygo homo sapien  |
| 11         | 33    | 100.0       | 419    | 16 Q7U5E3 | Q7u5e3 synchococc   |
| 12         | 33    | 100.0       | 423    | 13 Q98975 | Q98975 oncorhynch   |
| 13         | 33    | 100.0       | 529    | 16 Q9KYCA | Q9kyca streptomyce  |
| 14         | 33    | 100.0       | 560    | 2 Q8KQ21  | Q8kq21 pseudomonas  |
| 15         | 33    | 100.0       | 560    | 16 Q8ED23 | Q8ed23 pseudomonas  |
| 16         | 33    | 100.0       | 576    | 11 Q8C7C7 | Q8c7c7 mus musculus |

17 33 100.0 584 6 Q7YSG3  
18 33 100.0 608 5 Q95VB7  
19 33 100.0 608 11 Q8C7H3  
20 33 100.0 706 16 Q8FW79  
21 33 100.0 714 16 Q8YC46  
22 31 93.9 91 8 Q9B968  
23 31 93.9 132 17 Q9YFR4  
24 31 93.9 206 16 Q89CK9  
25 31 93.9 287 13 Q8JJ44  
26 31 93.9 314 16 Q892R2  
27 31 93.9 317 8 Q9TNJ6  
28 31 93.9 322 11 Q8C0J9  
29 31 93.9 323 11 Q8C1L9  
30 31 93.9 348 16 Q9RV18  
31 31 93.9 352 4 Q86W83  
32 31 93.9 394 5 O16677  
33 31 93.9 550 16 Q92ME9  
34 31 93.9 558 2 Q9R3W4  
35 31 93.9 559 2 Q9Z1Y1  
36 31 93.9 559 2 Q8VW58  
37 31 93.9 559 2 Q8KQ23  
38 31 93.9 559 2 Q93MW5  
39 31 93.9 559 2 Q7X5K2  
40 31 93.9 559 16 Q8ND25  
41 31 93.9 721 4 Q8NCK8  
42 31 93.9 721 4 Q8N2H2  
43 31 93.9 875 11 Q8B203  
44 31 93.9 889 11 Q8KIY2  
45 31 93.9 974 2 Q93PD2

## ALIGNMENTS

### RESULT 1

Q7TQM6 PRELIMINARY; PRT; 123 AA.

AC Q7TQM6;  
DT 01-OCT-2003 (TRENBLrel. 25, Created)  
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
DE Inward-rectifying potassium channel KCNJ10 (Fragment).  
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
OC Meriones.  
OX NCBF\_TaxID=10047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Wangemann P., White E.M., Albrecht B., Wu T., Maganti R.J.,  
RA Jabba S.V., Lee J.H., Everett L.A., Royaux I.E., Green E.D.,  
RA Marcus D.C.;  
RT "Loss of KCNJ10 expression abolishes endocochlear potential and causes deafness in Pendred's syndrome mouse model."  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY308017; AAF74570.1; -  
KW Ionic channel.

FT NON\_TER 1  
FT NON\_TER 123 123  
SQ SEQUENCE 123 AA; 13412 MW; EDD52P98P8908A4E CRC64;  
Query Match 100.0%; Score 33; DB 11; Length 123;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

### RESULT 2

Q8WMJ3

QY 1 HGDLLLE 6

DB 4 HGDLLLE 9



Q8WMJ3 PRELIMINARY; PRT; 129 AA.

AC Q8WMJ3; 1  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Protocadherin alpha 8 (Fragment).  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
NCBI\_TaxID=9544;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Medial basal hypothalamus;  
RA Brown A.E., Ojeda S.R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF424824; AAL40256.1; --  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR DR InterPro; IPR002126; Cadherin.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM00112; CA; 1.  
DR PROSITE; PS00232; CADHERIN\_1; 1.  
DR PROSITE; PS00268; CADHERIN\_2; 1.  
FT NON TER 1  
FT NON TER 129  
SQ SEQUENCE 129 AA; 14483 MW; 2D345B931CB6D0592 CRC64;

Query Match 100.0%; Score 33; DB 6; Length 129;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLLE 6  
Db 63 HGDLLLE 68  
|||||  
|

RESULT 3  
Q8WMH8 PRELIMINARY; PRT; 194 AA.

AC Q8WMH8; 1  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Protocadherin alpha 11 (Fragment).  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
NCBI\_TaxID=9544;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Medial basal hypothalamus;  
RA Mungenast A.E., Ojeda S.R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF424842; AAL40256.1; --  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005194; F:cell adhesion molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.  
DR Pfam; PF00028; cadherin; 1.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM00112; CA; 2.  
DR PROSITE; PS00232; CADHERIN\_1; 1.  
DR PROSITE; PS00268; CADHERIN\_2; 2.  
KW Calcium; Calcium-binding; Cell adhesion; Glycoprotein.  
FT NON TER 1  
FT NON TER 194  
SQ SEQUENCE 194 AA; 21808 MW; FB93F0B6A454D747 CRC64;

```

Query Match      100.0%; Score 33; DB 6; Length 194;
Best Local Similarity 100.0%; Pred.No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      43 HGDLLLE 48

RESULT 4
C19044 PRELIMINARY; PRT; 211 AA.
ID Q19044
AC O19044;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE K1r4.1 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxId=9986;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Japanese White;
RA Iehii M., Horio Y., Tada Y., Hibino H., Inanobe A., Ito M., Yamada M.,
PA Getow T., Uchiyama Y., Kurachi Y.;
RT "expression and clustered distribution of an inwardly rectifying
KT potassium channel, KAB-2/Kir4.1, on mammalian retinal Muller cell
RL membrane: Their regulation by insulin and laminin signals.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DDJB databases.
DR ENBL; AB005734; BAA21549.1; -
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0005242; F:inward rectifier potassium channel activity; IEA.
DR GO; GO:0005267; F:potassium channel activity; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR GO; GO:0001838; K-channel_IR.
DR InterPro; IPRO01838; K-channel_IR.
DR InterPro; IPRO01622; K-channel_pore.
DR Pfam; PF01007; IRK; 1.
DR PRINTS; PRO1320; KIRCHANNEL.
DR ProDom; PD001103; K+channel_IR; 2.
FT NON_TER 1
FT NON_TER 211
SQ SEQUENCE 211 AA; 23279 MW; 1BC61D158AD09E3D CRC64;

Query Match      100.0%; Score 33; DB 6; Length 211;
Best Local Similarity 100.0%; Pred.No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      22 HGDLLLE 27

RESULT 5
Q916L2 PRELIMINARY; PRT; 250 AA.
ID Q916L2
AC Q916L2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein PA0278.
DE Hypothetical protein PA0278.
GN PA0278.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OC NCBI_TaxId=287;
[1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.I., Huftnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

```

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
RT opportunistic pathogen."  
RL Nature 406:959-964 (2000).  
DR ENBL: AE004466; AAG03667.1; -.  
DR F83609; F83609. F83609. DUF81.  
DR InterPro: IPR002781; DUF81.  
DR Pfam: PF01925; DUF81; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 250 AA; 25773 MW; C5F5D079CEDA48A1 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 250;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
| | | | |  
Db 191 HGDLL 196

RESULT 6  
Q84712 PRELIMINARY; PRT; 251 AA.  
AC Q84712  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Pseudomonas putida.  
CG Plasmid pDTG1.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIB 9816-4;  
RA Zylstra G.J., Dennis J.J.;  
RT "Complete nucleotide sequence of the NAH plasmid pDTG1 from  
RT Pseudomonas putida NCIB 9816-4."  
RL Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.  
DR ENBL: AF491307; AA064230.1; -.  
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.  
DR InterPro: IPR006015; Usp.  
DR PRINTS: PR01438; UNYRS1STRESS.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 251 AA; 27295 MW; F9F4CEEFFB187FA5 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
| | | | |  
Db 73 HGDLL 78

RESULT 7  
Q9RBV8 PRELIMINARY; PRT; 283 AA.  
AC Q9RBV8  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Pseudomonas sp. R9.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=101164;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R9; TRANSPOSON=Tn1404;

RX MEDLINE=20011227; PubMed=10543801;  
RA Schnabel E.L., Jones A.L.;  
RT "Distribution of tetracycline resistance genes and transposons among  
RT phytophane bacteria in Michigan apple orchards."  
RL Appl. Environ. Microbiol. 65:4898-4907 (1999).  
DR ENBL: AF157797; AAD47991.1; -.  
DR GO: GO:0006950; P:response to stress; IEA.  
DR InterPro: IPR006016; Usp\_dom.  
DR Pfam: PF00582; Usp; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 283 AA; 30670 MW; 5586DD939254DDE CRC64;

Query Match 100.0%; Score 33; DB 2; Length 283;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
| | | | |  
Db 105 HGDLL 110

RESULT 8  
Q8N417 PRELIMINARY; PRT; 379 AA.  
AC Q8N417  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Potassium inwardly-rectifying channel, subfamily J, member 10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR ENBL: SC034036; AAH34036.1; -.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005242; P:inward rectifier potassium channel activity; IEA.  
DR GO: GO:0005267; P:potassium channel activity; IEA.  
DR GO: GO:0006813; P:potassium ion transport; IEA.  
DR InterPro: IPR001838; K-channel\_IR.  
DR InterPro: IPR001622; K-channel\_pore.  
DR Pfam: PF01007; IRK; 1.  
DR PRINTS: PR01320; KIRCHANNEL.  
DR ProDom: PD001103; K-channel\_IR; 2.  
SQ SEQUENCE 379 AA; 42550 MW; 9D8E156BEECE4CDD CRC64;

Query Match 100.0%; Score 33; DB 4; Length 379;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
| | | | |  
Db 93 HGDLL 98

RESULT 9  
Q8C7Z5 PRELIMINARY; PRT; 379 AA.  
AC Q8C7Z5  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Potassium inwardly-rectifying channel.  
GN KCNJ10.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
DR EMBL; AK048864; BAC33477.1; -  
DR MGD; NGI:1194504; Kcnj10.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005242; F:inward rectifier potassium channel activity; IEA.  
DR GO; GO:0005267; F:potassium channel activity; IEA.  
DR GO; GO:0006813; P:potassium ion transport; IEA.  
DR InterPro; IPR001838; K-channel\_IR.  
DR InterPro; IPR001622; K-channel\_pore.  
DR Pfam; PF01007; IRK; 1  
DR PRINTS; PR01320; KIRCHANNEL.  
DR PRODOM; PD001103; K+channel\_IR; 2.  
SQ SEQUENCE 379 AA; 42459 MW; 7FF6E1360F62E9B3 CRC64;  
  
Query Match 100.0%; Score 33; DB 11; Length 379;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 HGDLL 6  
Db 93 HGDLL 98  
  
RESULT 10  
Q86YGO PRELIMINARY; PRT; 417 AA.  
AC Q86YGO;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Similar to alpha-fetoprotein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Liver;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041789; AAH41789.1; -  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR00264; Serum albumin.  
DR Pfam; PF00273; transport\_Prot; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR PRODOM; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEFAE8D CRC64;  
  
Query Match 100.0%; Score 33; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 HGDLL 6  
Db 79 HGDLL 84  
  
RESULT 11  
Q7USE3 PRELIMINARY; PRT; 419 AA.  
ID Q7USE3;  
AC Q7USE3;  
DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Putative dihydroorotase (EC 3.5.2.3).  
GN SYNW1764.  
OS Synchococcus sp. (strain WH8102).  
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.  
OX NCBI\_TaxID=84588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=22825697; PubMed=12917641;  
RA Palenik B., Brahamsa B., Larimer F.W., Land M., Hauser L., Chain P.,  
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,  
RA Lufresne A., Patensky F., Webb E.A., Waterbury J.;  
RT "The genome of a motile marine Synchococcus";  
RL Nature 424:1037-1042(2003).  
DR EMBL; BX569693; CAE08279.1; -  
DR Hydrolase; Complete proteome.  
KW Hydrolase;  
SQ SEQUENCE 419 AA; 44483 MW; DFF0860393F3298B CRC64;  
  
Query Match 100.0%; Score 33; DB 16; Length 419;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 HGDLL 6  
Db 142 HGDLL 147  
  
RESULT 12  
Q98975 PRELIMINARY; PRT; 423 AA.  
AC Q98975;  
DT 01-FEB-1997 (TRENBLrel. 02, Created)  
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Weakly inward rectifying potassium channel.  
GN SWIRK.  
OS Oncorhynchus masou (Cherry salmon) (Masu salmon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8020;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96279099; PubMed=8663136;  
RA Kubo Y., Miyashita T., Kubokawa K.;  
RT "A weakly inward rectifying potassium channel of the salmon brain:  
RT Glutamate 179 in the second transmembrane domain is insufficient for  
RT strong rectification";  
RL J. Biol. Chem. 271:15729-15735(1996).  
DR EMBL; D83537; BAA12009.2; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005242; F:inward rectifier potassium channel activity; IEA.  
DR GO; GO:0005216; F:ion channel activity; IEA.  
DR GO; GO:0005267; P:potassium channel activity; IEA.  
DR GO; GO:0006813; P:potassium ion transport; IEA.  
DR InterPro; IPR001838; K-channel\_IR.  
DR InterPro; IPR001622; K-channel\_pore.  
DR Pfam; PF01007; IRK; 1.  
DR PRINTS; PR01320; KIRCHANNEL.  
DR PRODOM; PD001103; K+channel\_IR; 2.  
KW Ionic channel.  
SQ SEQUENCE 423 AA; 46950 MW; CDDF0159275EAE79 CRC64;  
  
Query Match 100.0%; Score 33; DB 13; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 HGDLL 6  
Db 111 HGDLL 116

## RESULT 13

Q9KYC4 PRELIMINARY; PRT; 529 AA.  
 AC Q9KYC4  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein SCO6906.  
 GN SCO6906 OR SC1B2.12C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL939129; CAB92568.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 529 AA; 57433 MW; 87075B0602262E11 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 529;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6  
 |||||  
 Db 182 HGDLLE 187

## RESULT 14

Q8KQ21 PRELIMINARY; PRT; 560 AA.  
 AC Q8KQ21  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE PHA synthase 2.  
 GN PHA2.  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lee Y.-H., Kim T.-K., Shin H.-D., Lee J.-N., Seo M.-C.;  
 RT "Molecular structure of PCR cloned PHA synthase I gene of Pseudomonas  
 putida KT2440 and its utilization for medium chain length  
 polyhydroxyalkanoate production.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY113181; AA63409.1; -.  
 DR GO; GO:0003993; Fracid phosphatase activity; IEA.  
 DR InterPro; IPR000073; A/b hydrolase.  
 DR InterPro; IPR000560; HisAc\_phsphtse.  
 DR Pfam; PF00561; abhydrolase; 1.  
 DR PROSITE; PS00616; HIS ACID\_PHOSPHAT\_1; 1.  
 SQ SEQUENCE 560 AA; 62640 MW; COA768C17569B4D3 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6  
 |||||  
 Db 410 HGDLLE 415

## RESULT 15

Q88D23 PRELIMINARY; PRT; 560 AA.  
 AC Q88D23;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Poly(3-hydroxyalkanoate) polymerase 2.  
 GN PHAC OR PP5005.  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L., Bauman M., DeBoy R.T., Daugherty S., Kolonay J.,  
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,  
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Eisen J., Timmis K.N., Duesternoeft A., Tuenmler B.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002).  
 DR EMBL; AF016792; AAN70571.1; -.  
 DR TIGR; PF000593; Fracid phosphatase activity; IEA.  
 DR GO; GO:0003993; Fracid phosphatase activity; IEA.  
 DR InterPro; IPR000073; A/b hydrolase.  
 DR InterPro; IPR000560; HisAc\_phsphtse.  
 DR Pfam; PF00561; abhydrolase; 1.  
 DR PROSITE; PS00616; HIS ACID\_PHOSPHAT\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 560 AA; 62722 MW; 4BF0D75DB5874868 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6  
 |||||  
 Db 410 HGDLLE 415

Search completed: April 19, 2004, 12:00:07  
 Job time : 4.31579 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 3.52909 Seconds  
(without alignments)

480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1980s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description |
|------------|-------|-------------|--------|----------|-------------|
| 1          | 33    | 100.0       | 14     | ABB52153 | Human API   |
| 2          | 33    | 100.0       | 14     | ABP61551 | Human KRP   |
| 3          | 33    | 100.0       | 14     | ABR58983 | Alzheimer   |
| 4          | 33    | 100.0       | 14     | ABR75499 | Liver res   |
| 5          | 33    | 100.0       | 18     | ABR12469 | Human alb   |
| 6          | 33    | 100.0       | 21     | ABR62326 | Bovine se   |
| 7          | 33    | 100.0       | 58     | ABB40958 | Peptide #   |
| 8          | 33    | 100.0       | 58     | AAW34732 | Peptide #   |
| 9          | 33    | 100.0       | 58     | ABB25068 | Protein #   |
| 10         | 33    | 100.0       | 58     | AAW74619 | Human bon   |
| 11         | 33    | 100.0       | 58     | AAW61818 | Human bra   |
| 12         | 33    | 100.0       | 58     | ABG56405 | Human liv   |
| 13         | 33    | 100.0       | 58     | ABG44448 | Human pep   |
| 14         | 33    | 100.0       | 85     | AAW18788 | Peptide #   |
| 15         | 33    | 100.0       | 85     | ABB37889 | Peptide #   |
| 16         | 33    | 100.0       | 85     | AAW31296 | Peptide #   |
| 17         | 33    | 100.0       | 85     | ABB23147 | Protein #   |
| 18         | 33    | 100.0       | 85     | AAW71019 | Human bon   |
| 19         | 33    | 100.0       | 85     | AAW58521 | Human bra   |
| 20         | 33    | 100.0       | 85     | ABG52734 | Human liv   |
| 21         | 33    | 100.0       | 85     | ABG40810 | Human pep   |
| 22         | 33    | 100.0       | 114    | AAW25284 | Human pro   |
| 23         | 33    | 100.0       | 116    | AAW00108 | Human pol   |
| 24         | 33    | 100.0       | 123    | AAW23861 | abCEST en   |
| 25         | 33    | 100.0       | 126    | AAW04435 | Human pol   |

|    |    |       |     |   |          |           |
|----|----|-------|-----|---|----------|-----------|
| 26 | 33 | 100.0 | 131 | 4 | ABG01238 | Novel hum |
| 27 | 33 | 100.0 | 143 | 4 | AAO02571 | Human pol |
| 28 | 33 | 100.0 | 188 | 3 | AAW83948 | Yeast cod |
| 29 | 33 | 100.0 | 265 | 7 | AAW25672 | Dog aller |
| 30 | 33 | 100.0 | 265 | 7 | ADC34920 | Dog aller |
| 31 | 33 | 100.0 | 293 | 4 | AAW33082 | Novel hum |
| 32 | 33 | 100.0 | 303 | 2 | AAW14178 | Human ser |
| 33 | 33 | 100.0 | 365 | 5 | ABP51353 | Human MDD |
| 34 | 33 | 100.0 | 373 | 1 | AAW90387 | N-termina |
| 35 | 33 | 100.0 | 373 | 1 | AAW90387 | N-termina |
| 36 | 33 | 100.0 | 379 | 2 | AAW26368 | Human kid |
| 37 | 33 | 100.0 | 379 | 5 | AAW47972 | Human KIR |
| 38 | 33 | 100.0 | 379 | 5 | AAW47972 | Human KIR |
| 39 | 33 | 100.0 | 388 | 1 | AAW90389 | N-termina |
| 40 | 33 | 100.0 | 389 | 1 | AAW90390 | N-termina |
| 41 | 33 | 100.0 | 390 | 1 | AAW90391 | N-termina |
| 42 | 33 | 100.0 | 407 | 1 | AAW90392 | N-termina |
| 43 | 33 | 100.0 | 463 | 2 | AAW14179 | Human ser |
| 44 | 33 | 100.0 | 500 | 7 | ADD32019 | Heterolog |
| 45 | 33 | 100.0 | 550 | 4 | AAW29877 | Novel hum |

#### ALIGNMENTS

#### RESULT 1

ABB52153

ID ABB52153 standard; peptide; 14 AA.

XX

AC ABB52153;

DT 08-FEB-2002 (first entry)

DE Human API-95 tryptic digest peptide #1.

XX

KW Human; neuroprotective; nootropic; gene therapy; vaccine;

KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;

KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;

KW Expression Reference Protein Isoform; ERPI; proteolysis.

XX

OS Homo sapiens.

XX

PN WO200175454-A2.

XX

PD 11-OCT-2001.

XX

PF 03-APR-2001; 2001WO-US010908.

XX

PR 03-APR-2000; 2000US-0194504P.

PR 28-NOV-2000; 2000US-0253647P.

XX

PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

PA (PFIZ ) PFIZER INC.

XX

PI Durham KL, Friedman DL, Herath HM, Kimmel LH, Parekh RB;

PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;

PI Townsend RR, White F, Williams SA;

XX

DR WPI; 2001-639384/73.

XX

PT Screening for Alzheimer's disease in a mammal, by making two-dimensional

PT array of a feature whose relative abundance correlates with disease, and

PT comparing with abundance of the feature in samples of healthy persons.

XX

PS Example; Page 29; 162pp; English.

XX

CC The invention relates to methods for the screening, diagnosis and

CC prognosis of Alzheimer's disease. The methods involve the detection of

CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-

CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or

CC plasma. The abundance of the AFs and APIs is then normalised to an

CC Expression Reference Protein Isoform (ERPI) in order to determine whether

CC a patient is suffering from, or has a predisposition to, Alzheimer's

CC Disease. The relative abundance of the Afs and APIs correlates with the  
 CC severity of Alzheimer's Disease. The present sequence is a peptide  
 CC produced from an API by proteolysis  
 XX  
 SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6  
 DB 4 HGDLLLE 9

RESULT 2  
 ABP61551  
 ID ABP61551 standard; peptide; 14 AA.  
 XX  
 AC ABP61551;  
 XX  
 DT 02-OCT-2002 (first entry)  
 XX  
 DE Human KRPI tryptic digest peptide #46.  
 XX  
 KW Human; tryptic digest peptide; KRPI; kidney response; KR;  
 KW nephrotropic kidney response-associated protein isoform; gene therapy;  
 KW antitense therapy; kidney function; tubular nephritis; renal failure;  
 KW nephron cell metabolic pathway modulation; glomerular necrosis;  
 KW papillary necrosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200254081-A2.  
 XX  
 PD 11-JUL-2002.  
 XX  
 XX 24-DEC-2001; 2001WO-GB005777.  
 XX  
 PR 29-DEC-2000; 2000US-0260392P.  
 XX  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 PI Holt GD, Kelly MD, Kennedy SJ, Moyses C;  
 XX  
 DR WPI; 2002-583637/62.  
 XX  
 XX Screening, diagnosis or prognosis of kidney response in subject, by  
 PT detecting kidney response-associated features or kidney response-  
 PT associated protein isoforms in body fluid or tissue from subject.  
 XX  
 PS Disclosure; Page 40; 168pp; English.  
 XX  
 CC The invention relates to a novel method for the screening, diagnosis or  
 CC prognosis of kidney response (KR). The method of the invention has  
 CC nephrotropic activity, and may have a use in gene therapy or antisense  
 CC therapy. The method is useful for the screening, diagnosis or prognosis  
 CC of KR in a subject, for determining the stage or severity of KR in a  
 CC subject, for identifying a subject at risk of developing KR, or for  
 CC monitoring the effect of therapy administered to a subject with KR. An  
 CC alternative method of the invention is useful for screening agents that  
 CC interact with one or more of the kidney response-associated protein  
 CC isoforms (KRPIs). The kidney response includes alterations in kidney  
 CC function, any phase of nephron cell metabolic pathway modulation,  
 CC glomerular/proximal tubular nephritis, glomerular/papillary necrosis,  
 CC acute and chronic renal failure, and end stage renal disease. The  
 CC sequences shown in ABP61514-ABP61787 represent tryptic digest peptides of  
 CC the KRPIs of the invention.  
 XX  
 SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.5;

OY 1 HGDLLLE 6  
 DB 4 HGDLLLE 9

RESULT 4  
 ABR75499  
 ID ABR75499 standard; peptide; 14 AA.  
 XX  
 AC ABR75499;  
 XX  
 DT 28-AUG-2003 (first entry)

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6  
 DB 4 HGDLLLE 9

RESULT 3  
 ABR58983  
 ID ABR58983 standard; peptide; 14 AA.  
 XX  
 AC ABR58983;  
 XX  
 DT 11-JUL-2003 (first entry)  
 XX  
 DE Alzheimer's Disease-associated protein isoform, API-95, SEQ ID 77.  
 XX  
 KW Nootropic; Neuroprotective; Alzheimer's disease; API; human;  
 KW Alzheimer's Disease-associated protein isoform.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003028543-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 XX 03-OCT-2002; 2002WO-US031642.  
 PF  
 XX 03-OCT-2001; 2001US-0326708P.  
 PR  
 XX (PFIZ ) PFIZER PROD INC.  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 XX Durham IK, Friedman DL, Herath HM, Kimmel LH, Parekh RB;  
 PI Potter DM, Rohlf C, Silber BM, Snyder PJ, Soares HD, Stiger TR;  
 PI Sunderland PT, Townsend RR, White WF, Williams SA;  
 XX  
 DR WPI; 2003-371957/35.  
 XX  
 PT Screening or diagnosing of Alzheimer's disease (AD) determine the stage  
 PT or severity of AD in a subject, comprises analyzing a test sample of body  
 PT fluid from the subject by 2-dimensional electrophoresis.  
 XX  
 PS Claim 2; Page 43; 179pp; English.  
 XX  
 CC The present invention relates to methods for screening or diagnosing  
 CC Alzheimer's disease (AD) to determine the stage or severity of AD in a  
 CC subject to identify subject at risk of developing AD, or to monitor the  
 CC effect of therapy administered. The methods comprise analysing a test  
 CC sample of body fluid by 2-dimensional electrophoresis to generate a 2-  
 CC dimensional array of AD-associated features (AFs). The method  
 CC alternatively comprises quantitatively detecting in a sample of body  
 CC fluid from the subject, one or more AD-associated protein isoforms (APIs;  
 CC ABR58710-ABR59184)  
 XX  
 SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 6; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6  
 DB 4 HGDLLLE 9

RESULT 4  
 ABR75499  
 ID ABR75499 standard; peptide; 14 AA.  
 XX  
 AC ABR75499;  
 XX  
 DT 28-AUG-2003 (first entry)

XX Liver response-associated protein isoform (LRPI) peptide SEQ ID NO:84.  
DE Biomarker; liver response; liver response-associated protein isoform;  
XX LRPI; liver response-associated feature; LRF.  
KW Synthetic.  
OS WO2003038444-A2.  
XX 08-MAY-2003.  
XX 31-OCT-2002; 2002WO-US034847.  
XX 31-OCT-2001; 2001US-0335964P.  
XX (PFIZ ) PFIZER PROD INC.  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX Amacher DE, Fasulo LM, Herath HMAc, Holt GD, Stiger TR;  
XX WPI; 2003-430566/40.  
XX Screening, diagnosing, staging or identifying subject at risk of  
XX developing, liver response, or monitoring effect of therapy on liver  
XX response, by detecting Liver Response-Associated Protein Isoforms in  
XX subject sample.  
XX Claim 1; Page 55; 256pp; English.  
XX The present invention describes a method (M1) for screening or diagnosing  
XX a liver response in a subject, determining the stage or severity of a  
XX liver response in a subject, identifying a subject at risk of developing  
XX liver response, or monitoring the effect of therapy administered to a  
XX subject having liver response, involving detecting liver response-  
XX associated protein isoforms (LRPIs) (see the peptides given in ABR75395  
XX to ABR75806) in a test biological sample from the subject. Alternatively,  
XX screening or diagnosing a liver response in a subject, or monitoring the  
XX effect of a drug or therapy administered to a subject, involves  
XX contacting at least one oligonucleotide probe comprising 10 or more  
XX consecutive nucleotides complementary to a nucleotide sequence encoding  
XX an LRPI with RNA obtained from a biological sample from the subject or  
XX with cDNA copied from the RNA, where the contacting occurs under  
XX conditions that permit hybridisation of the probe to the nucleotide  
XX sequence if present, detecting hybridisation, if any, between the probe  
XX and the nucleotide sequence, and comparing the hybridisation, if any,  
XX detected in the above step, with the hybridisation detected in a control  
XX sample, or with a previously determined reference range. M1 is useful for  
XX screening or diagnosing a liver response in a subject, determining the  
XX stage or severity of a liver response in a subject, identifying a subject  
XX at risk of developing liver response, and monitoring the effect of  
XX therapy administered to a subject having liver response  
XX Sequence 14 AA;  
SQ  
Query Match 100.0%; Score 33; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
DB 4 HGDLLLE 9  
RESULT 5  
ID AAB12469 standard; peptide; 18 AA.  
XX AAB12469;  
AC AAB12469;  
XX 25-OCT-2000 (first entry)  
DT Human albumin epitope peptide #8.  
XX Processing of protein or peptide samples used in mass spectrometer  
DE

XX Monoclonal antibody; hybridoma cell; immunoglobulin; IgG; fusion;  
KW human albumin; diabetic nephropathy; diagnosis.  
XX Homo sapiens.  
OS JP2000139460-A.  
XX 23-MAY-2000.  
XX 02-NOV-1998; 98JP-00311677.  
XX 02-NOV-1998; 98JP-00311677.  
XX (TOYW ) TOYOTA CHUO KENKYUSHO KK.  
XX (AISE ) AISIN SEIKI KK.  
XX WPI; 2000-433935/38.  
XX Hybridoma cells for preparation of IgG monoclonal antibody capable of  
XX rapid reaction with human albumin for diagnosis of diabetic nephropathy.  
XX Example; Fig 1; ilpp; Japanese.  
XX The present invention describes hybridoma cells used for preparing an  
XX immunoglobulin G (IgG) monoclonal antibody capable of rapid reaction with  
XX human albumin. The hybridoma cells are prepared by fusion of mammal  
XX myeloma cells and spleen cells immunised with human serum albumin, and  
XX producing IgG monoclonal antibody which rapidly react with human albumin.  
XX The monoclonal antibodies can be used in the diagnosis of diabetic  
XX nephropathy. The present sequence represents a human albumin epitope  
XX peptide sequence which is used in an example from the present invention  
XX Sequence 18 AA;  
SQ  
Query Match 100.0%; Score 33; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
DB 1 HGDLLLE 6  
RESULT 6  
ID ABR62926 standard; peptide; 21 AA.  
XX ABR62926  
XX ABR62926;  
AC ABR62926;  
XX 04-DEC-2003 (first entry)  
DT Bovine serum albumin peptide, analysed by mass spectrometry.  
XX DE Serum albumin; cattle; mass spectrometry; proteomics.  
KW Serum albumin; cattle; mass spectrometry; proteomics.  
XX Bos taurus.  
XX WO2003060524-A2.  
XX 24-JUL-2003.  
XX 30-DEC-2002; 2002WO-CA002024.  
XX 28-DEC-2001; 2001US-0343859P.  
XX (MDS-) MDS PROTEOMICS INC.  
XX Duwel H, Goh T, Lebihan T;  
XX WPI; 2003-618204/58.  
XX Processing of protein or peptide samples used in mass spectrometer  
PT

PT analysis, by reversibly immobilizing samples onto solid support,  
PT transforming to solid-phase chemical unit, eluting and recovering  
PT fragments from support.  
XX  
PS Example 4; Page 46; 66pp; English.  
XX  
CC The present sequence is a peptide fragment of bovine serum albumin (BSA).  
CC The peptide was produced by chemical reduction of cysteine residues with  
CC dithiothreitol, alkylation with iodoacetamide and enzymatic digestion  
CC with trypsin. This is an example of the protein processing method of the  
CC invention, which was designed for use in proteomics analysis by mass  
CC spectrometry (MS). The general method involves: reversibly immobilising  
CC protein or peptide samples onto a solid support, e.g. a cation-exchange  
CC resin; subjecting the immobilised protein or peptide to solid phase-based  
CC chemical and/or enzymatic modifications; eluting the resulting peptide or  
CC protein derived fragments from the solid support; and recovering the  
CC peptides in near-quantitative yields in a format suitable for  
CC identification by routine MS technologies  
XX  
SQ Sequence 21 AA;  
  
Query Match 100.0%; Score 33; DB 7; Length 21;  
Best Local Similarity 100.0%; Pred. No. 8.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HGDLLLE 6  
DB 4 HGDLLLE 9  
  
RESULT 7  
ABE40958  
ID ABB40958 standard; peptide; 58 AA.  
XX  
AC ABB40958;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #8464 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000669.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human foetal liver.  
XX  
PS Claim 27; SEQ ID NO 33593; 639pp + Sequence Listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 58 AA;  
  
Query Match 100.0%; Score 33; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HGDLLLE 6  
DB 34 HGDLLLE 39  
  
RESULT 8  
AAM34732  
ID AAM34732 standard; protein; 58 AA.  
XX  
AC AAM34732;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #8769 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000663.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488997/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
PS Claim 27; SEQ ID NO 35001; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SNP:  
CC see AAI31315-AAI57546). the present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders  
XX  
SQ Sequence 58 AA;  
  
Query Match 100.0%; Score 33; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HGDLLLE 6  
DB 34 HGDLLLE 39



```

RESULT 9
ABB25068
ID ABB25068 standard; protein; 58 AA.
XX AC ABB25068;
XX DT 23-JAN-2002 (first entry)
XX DE Protein #7067 encoded by probe for measuring heart cell gene expression.
XX DE Human; gene expression; heart; microarray; vascular system;
XX DE cardiovascular disease; hypertension; cardiac arrhythmia;
XX DE congenital heart disease.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234887P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488990/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 15; SEQ ID NO 26838; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart (see
XX CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX CC probe. The probes may be used for predicting, measuring and displaying
XX CC gene expression in samples derived from the human heart via microarrays.
XX CC By measuring gene expression, the probes are useful for predicting,
XX CC diagnosing, grading, staging, monitoring and prognosing diseases of the
XX CC human heart and vascular system e.g. cardiovascular disease, Note: The
XX CC hypertension, cardiac arrhythmias and congenital heart disease. The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 58 AA;
XX Query Match 100.0%; Score 33; DB 4; Length 58;
XX Best Local Similarity 100.0%; Pred. No. 22;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 HGDLLLE 6
XX DB 34 HGDLLLE 39
XX RESULT 10
XX AAM74619
XX ID AAM74619 standard; protein; 58 AA.
XX AC AAM74619;
XX DT 06-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33923.
XX DE Human; brain expressed exon; gene expression analysis; probe; microarray;
XX DE Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX
```

```

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34925.
XX DE Human; bone marrow expressed exon; gene expression analysis; probe;
XX DE microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234887P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488990/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 34925; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX SQ Sequence 58 AA;
XX Query Match 100.0%; Score 33; DB 4; Length 58;
XX Best Local Similarity 100.0%; Pred. No. 22;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 HGDLLLE 6
XX DB 34 HGDLLLE 39
XX RESULT 11
XX AAM61818
XX ID AAM61818 standard; protein; 58 AA.
XX AC AAM61818;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33923.
XX DE Human; brain expressed exon; gene expression analysis; probe; microarray;
XX DE Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX
```

PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483446/52.  
XX Single exon nucleic acid probes for analyzing gene expression in human  
XX brains.  
XX Example 4; SEQ ID NO 33923; 650pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention  
XX  
XX Sequence 58 AA;  
Query Match 100.0%; Score 33; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
Db 34 HGDLLLE 39  
RESULT 12  
ABG56405  
ID ABG56405 standard; peptide; 58 AA.  
XX AC ABG56405;  
XX DT 25-FEB-2003 (first entry)  
XX DE Human liver peptide, SEQ ID NO 35053.  
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
XX hypercholesterolaemia; coronary heart disease.  
XX OS Homo sapiens.  
XX PN WO200157273-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000664.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488898/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human adult liver.  
XX Claim 27; SEQ ID NO 35053; 658pp; English.  
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
XX measuring human gene expression in a sample derived from human adult  
XX liver, comprising one of 13109 defined nucleotide sequences given in the  
XX specification (or complements/ fragments). The probe hybridises at high  
XX stringency to a nucleic acid molecule expressed in the human adult liver.  
XX (I) may be used for predicting, measuring and displaying gene expression  
XX in samples derived from human adult liver. The genes identified may be  
XX involved in genetic liver diseases such as cirrhosis,  
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
XX associated with coronary heart disease. ABG47348-ABG59930 represent human  
XX liver single exon encoded peptides of the invention. Note: the sequence  
XX information for this patent does not appear in the printed specification  
XX but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 58 AA;  
Query Match 100.0%; Score 33; DB 4; Length 58;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
Db 34 HGDLLLE 39  
RESULT 13  
ABG44448  
ID ABG44448 standard; peptide; 58 AA.  
XX AC ABG44448;  
XX DT 19-AUG-2002 (first entry)  
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 34113.  
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemostodiosis;  
XX pulmonary histiocytosis; lymphangioma; lymphoma; Karagener syndrome;  
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
XX OS Homo sapiens.  
XX PN WO200186003-A2.  
XX PD 15-NOV-2001.  
XX PF 30-JAN-2001; 2001WO-US000665.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.  
XX Claim 27; SEQ ID NO 34113; 634pp; English.  
PS  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12347 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC ; the novel set of probes which hybridise at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a peptide/protein encoded by a single exon probe of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 58 AA;  
S

Query Match 100.0%; Score 33; DB 5; Length 58;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 HGDLLLE 6  
Db 34 HGDLLLE 39  
RESULT 14  
AAM18788  
ID AAM18788 standard; protein; 85 AA.  
AC  
XX AAM18788;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Peptide #5222 encoded by probe for measuring cervical gene expression.  
XX Probe; human; microarray; gene expression; cervical epithelial cell;  
KW Cervical cancer.  
XX  
XX Homo sapiens.  
OS  
XX WO200157278-A2.  
PN  
XX  
PD 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US000670.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human cervical epithelial cells.  
PT Claim 27; SEQ ID NO 23614; 487pp; English.  
PS  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 85 AA;  
S

Query Match 100.0%; Score 33; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 HGDLLLE 6  
Db 34 HGDLLLE 39  
RESULT 15  
ABB37889  
ID ABB37889 standard; peptide; 85 AA.  
XX  
AC ABB37889;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #5395 encoded by human foetal liver single exon probe.  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX Homo sapiens.  
OS  
XX WO200157277-A2.  
PN  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000669.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX

```

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-493447/52.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT Gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 30524; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC single exon expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 85 AA;
Query Match 100.0%; Score 33; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLLE 6
Db 34 HGDLLLE 39
Search completed: April 19, 2004, 11:51:19
Job time : 4.52909 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.52632 Seconds

(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLL6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 33    | 100.0       | 14     | 9  | US-09-826-290-168   |
| 2          | 33    | 100.0       | 14     | 12 | US-10-601-837-46    |
| 3          | 33    | 100.0       | 14     | 15 | US-10-285-394-84    |
| 4          | 33    | 100.0       | 14     | 16 | US-10-264-309-77    |
| 5          | 33    | 100.0       | 58     | 9  | US-09-864-761-40366 |
| 6          | 33    | 100.0       | 85     | 9  | US-09-864-761-38445 |
| 7          | 33    | 100.0       | 99     | 14 | US-10-029-386-29821 |
| 8          | 33    | 100.0       | 111    | 14 | US-10-029-386-29303 |
| 9          | 33    | 100.0       | 111    | 14 | US-10-029-386-29468 |
| 10         | 33    | 100.0       | 114    | 12 | US-10-296-115-799   |
| 11         | 33    | 100.0       | 114    | 14 | US-10-029-386-29057 |
| 12         | 33    | 100.0       | 119    | 14 | US-10-029-386-29571 |
| 13         | 33    | 100.0       | 126    | 14 | US-10-029-386-29777 |
| 14         | 33    | 100.0       | 244    | 12 | US-10-425-114-57360 |
| 15         | 33    | 100.0       | 379    | 15 | US-10-353-690-88    |

|    |    |       |     |    |                     |                   |
|----|----|-------|-----|----|---------------------|-------------------|
| 16 | 33 | 100.0 | 379 | 15 | US-10-344-890-2     | Sequence 2, Appli |
| 17 | 33 | 100.0 | 553 | 12 | US-10-425-114-64794 | Sequence 64794, A |
| 18 | 33 | 100.0 | 583 | 15 | US-10-360-101-200   | Sequence 200, App |
| 19 | 33 | 100.0 | 585 | 9  | US-09-929-552-2     | Sequence 2, Appli |
| 20 | 33 | 100.0 | 585 | 10 | US-09-932-613-445   | Sequence 445, App |
| 21 | 33 | 100.0 | 585 | 10 | US-09-984-010-26    | Sequence 26, Appl |
| 22 | 33 | 100.0 | 585 | 10 | US-09-833-041-18    | Sequence 18, Appl |
| 23 | 33 | 100.0 | 585 | 10 | US-09-833-117-18    | Sequence 18, Appl |
| 24 | 33 | 100.0 | 585 | 10 | US-09-932-322-445   | Sequence 445, App |
| 25 | 33 | 100.0 | 585 | 10 | US-09-832-501-18    | Sequence 18, Appl |
| 26 | 33 | 100.0 | 585 | 11 | US-09-833-118-18    | Sequence 18, Appl |
| 27 | 33 | 100.0 | 585 | 11 | US-09-833-245-18    | Sequence 18, Appl |
| 28 | 33 | 100.0 | 585 | 12 | US-10-424-999-11    | Sequence 31, Appl |
| 29 | 33 | 100.0 | 585 | 12 | US-10-425-000-31    | Sequence 34, Appl |
| 30 | 33 | 100.0 | 585 | 12 | US-10-433-108-34    | Sequence 5, Appli |
| 31 | 33 | 100.0 | 585 | 13 | US-10-153-064-5     | Sequence 5, Appli |
| 32 | 33 | 100.0 | 585 | 14 | US-10-153-6048-5    | Sequence 1, Appli |
| 33 | 33 | 100.0 | 585 | 14 | US-10-319-263-1     | Sequence 2, Appli |
| 34 | 33 | 100.0 | 585 | 14 | US-10-319-263-2     | Sequence 1, Appli |
| 35 | 33 | 100.0 | 585 | 14 | US-10-414-469-1     | Sequence 1, Appli |
| 36 | 33 | 100.0 | 585 | 14 | US-10-414-469-2     | Sequence 2, Appli |
| 37 | 33 | 100.0 | 585 | 14 | US-10-413-831-1     | Sequence 1, Appli |
| 38 | 33 | 100.0 | 585 | 14 | US-10-413-831-2     | Sequence 1, Appli |
| 39 | 33 | 100.0 | 585 | 15 | US-10-413-832-1     | Sequence 2, Appli |
| 40 | 33 | 100.0 | 585 | 15 | US-10-413-832-2     | Sequence 1, Appli |
| 41 | 33 | 100.0 | 585 | 15 | US-10-414-386-1     | Sequence 2, Appli |
| 42 | 33 | 100.0 | 585 | 15 | US-10-414-386-2     | Sequence 11, Appl |
| 43 | 33 | 100.0 | 585 | 15 | US-10-233-675A-11   | Sequence 26, Appl |
| 44 | 33 | 100.0 | 585 | 15 | US-10-462-262-26    | Sequence 7, Appli |
| 45 | 33 | 100.0 | 604 | 10 | US-09-984-010-7     |                   |

#### ALIGNMENTS

#### RESULT 1

US-09-826-290-168  
; Sequence 168, Application US/09826290  
; Patent No. US2002016468A1  
; GENERAL INFORMATION:  
; APPLICANT: Durham, L. Kathryn  
; APPLICANT: Friedman, David L.  
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri  
; APPLICANT: Kimmel, Lida H.  
; APPLICANT: Patek, Rajesh Bhikhu  
; APPLICANT: Potter, David M.  
; APPLICANT: Rohlf, Christian  
; APPLICANT: Silber, B. Michael  
; APPLICANT: Stiger, Thomas R.  
; APPLICANT: Sunderland, P. Trey  
; APPLICANT: Townsend, Robert Reid  
; APPLICANT: White, Frost  
; APPLICANT: Williams, Stephen A.  
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and  
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of  
; FILE OF INVENTION: Alzheimer's Disease  
; FILE REFERENCE: 2572-1-001 N2  
; CURRENT APPLICATION NUMBER: US/09/826,290  
; CURRENT FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: US 60/194,504  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: US 60/253,647  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 492  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 168  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-09-826-290-168

Query Match 100.0%; Score 33; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
| | | | |  
Db 4 HGDLLLE 9

## RESULT 2

US-10-601-837-46  
; Sequence 46, Application US/10601837  
; Publication No. US2004005309A1  
; GENERAL INFORMATION:  
; APPLICANT: Holt, Gordon D  
; APPLICANT: Kelly, Michael D  
; APPLICANT: Kennedy, Sandra J  
; APPLICANT: Moyses, Christopher  
; TITLE OF INVENTION: Proteins, Genes and Their Use for Diagnosis and Treatment of Kidney Disease  
; FILE REFERENCE: 2543-1-030  
; CURRENT APPLICATION NUMBER: US/10/601.837  
; PRIOR FILING DATE: 2003-06-23  
; PRIOR APPLICATION NUMBER: PCT/GB01/05777  
; PRIOR FILING DATE: 2001-12-24  
; PRIOR APPLICATION NUMBER: US 60/260392  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 272  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Ratus No. US2004005309Alvegicus  
US-10-601-837-46

Query Match 100.0%; Score 33; DB 12; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
| | | | |  
Db 4 HGDLLLE 9

## RESULT 3

US-10-285-394-84  
; Sequence 84, Application US/10285394  
; Publication No. US20030228583A1  
; GENERAL INFORMATION:  
; APPLICANT: PASULO, LISA M.  
; APPLICANT: AMACHER, DAVID E.  
; APPLICANT: HERATH, HERATH MUDIYANSELAGE ATHULA CHANDRASIRI  
; APPLICANT: HOLT, GORDON DUANE  
; APPLICANT: STIGER, THOMAS R.  
; TITLE OF INVENTION: BIOMARKERS OF LIVER RESPONSE  
; FILE REFERENCE: POA-003.01  
; CURRENT APPLICATION NUMBER: US/10/285,394  
; PRIOR FILING DATE: 2003-02-07  
; PRIOR APPLICATION NUMBER: 60/335,964  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 412  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 84  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-285-394-84

Query Match 100.0%; Score 33; DB 15; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
| | | | |  
Db 4 HGDLLLE 9

## RESULT 4

US-10-264-309-77  
; Sequence 77, Application US/10264309  
; Publication No. US20040022794A1  
; GENERAL INFORMATION:  
; APPLICANT: DURHAM, L. KATHRYN  
; APPLICANT: FRIEDMAN, DAVID L.  
; APPLICANT: HERATH, HERATH  
; APPLICANT: KIMMEL, LIDA H.  
; APPLICANT: PAREKH, RAJESH B.  
; APPLICANT: POTTER, DAVID M.  
; APPLICANT: ROHLFF, CHRISTIAN  
; APPLICANT: SILBER, B. MICHAEL  
; APPLICANT: SNYDER, PETER J.  
; APPLICANT: SOARES, HOLLY D.  
; APPLICANT: STIGER, THOMAS R.  
; APPLICANT: SUNDERLAND, P. TREY  
; APPLICANT: TOMSEN, ROBERT R.  
; APPLICANT: WHITE, W. FROST  
; APPLICANT: WILLIAMS, STEPHEN A.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR.  
; FILE REFERENCE: POA-002.01  
; CURRENT APPLICATION NUMBER: US/10/264,309  
; PRIOR FILING DATE: 2002-10-03  
; PRIOR APPLICATION NUMBER: 60/326,708  
; PRIOR FILING DATE: 2001-10-03  
; NUMBER OF SEQ ID NOS: 491  
; SOFTWARE: PatentIn version 2.1  
; SEQ ID NO 77  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-309-77

Query Match 100.0%; Score 33; DB 16; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
| | | | |  
Db 4 HGDLLLE 9

## RESULT 5

US-09-864-761-40366  
; Sequence 40366, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmics-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 40366  
 LENGTH: 58  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC008468.4  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.89  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
 OTHER INFORMATION: EST HUMAN HIT: A196295.1, EVALUE 2.00e-25  
 OTHER INFORMATION: SWISSPROT HIT: O35865, EVALUE 4.60e+00  
 US-09-864-761-40366

Query Match 100.0%; Score 33; DB 9; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6  
 Db 34 HGDLL 39

RESULT 6  
 US-09-864-761-38445  
 Sequence 38445, Application US/09864761  
 Patent No. US20020048763A1  
 GENERAL INFORMATION:  
 APPLICANT: Rank, David R.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: Aecmca-X-1  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 38445  
 LENGTH: 85  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC005609.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
 OTHER INFORMATION: SWISSPROT HIT: O35865, EVALUE 8.30e+00  
 OTHER INFORMATION: EST HUMAN HIT: A196295.1, EVALUE 4.00e-41  
 US-09-864-761-38445

Query Match 100.0%; Score 33; DB 9; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6  
 Db 34 HGDLL 39

RESULT 7  
 US-10-029-386-29821  
 Sequence 29821, Application US/10029386  
 Publication No. US20030194704A1  
 GENERAL INFORMATION:  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: AECMCA-X-2  
 CURRENT FILING DATE: 2001-12-20  
 CURRENT APPLICATION NUMBER: US/10/029,386  
 NUMBER OF SEQ ID NOS: 34288  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 29821

; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHRS.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.94  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y511, EVALUATION 5.00e-51  
US-10-029-386-29468

Query Match 100.0%; Score 33; DB 14; Length 99;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||  
Db 42 HGDLLLE 47

## RESULT 8

US-10-029-386-29303  
; Sequence 29303, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29303  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHRS.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.65  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.49  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.41  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.47  
; OTHER INFORMATION: SWISSPROT HIT: Q9UN75, EVALUATION 1.00e-58  
US-10-029-386-29303

Query Match 100.0%; Score 33; DB 14; Length 111;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||  
Db 53 HGDLLLE 58

## RESULT 9

US-10-029-386-29468  
; Sequence 29468, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
US-10-029-386-29468

; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29468  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHRS.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.55  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y5H6, EVALUATION 2.00e-57  
US-10-029-386-29468

Query Match 100.0%; Score 33; DB 14; Length 111;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||  
Db 52 HGDLLLE 57

## RESULT 10

US-10-296-115-799  
; Sequence 799, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478  
; SEQ ID NO 799  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(114)  
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3  
US-10-296-115-799

Query Match 100.0%; Score 33; DB 12; Length 114;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||  
Db 46 HGDLLLE 51

## RESULT 11

US-10-029-386-29057  
; Sequence 29057, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
US-10-029-386-29057



; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29057  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR5.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.78  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y511, EVALUE 6.00e-60  
; US-10-029-386-29057

Query Match 100.0%; Score 33; DB 14; Length 114;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
DB 45 HGDLLLE 50

RESULT 12  
US-10-029-386-29571  
; Sequence 29571, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029.386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29571  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR5.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.71  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y512, EVALUE 1.00e-58  
; US-10-029-386-29571

Query Match 100.0%; Score 33; DB 14; Length 119;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
DB 60 HGDLLLE 65

RESULT 13  
US-10-029-386-29777  
; Sequence 29777, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029.386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29777  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR5.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.82  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.67  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: SWISSPROT HIT: Q9UN75, EVALUE 6.00e-63  
; US-10-029-386-29777

Query Match 100.0%; Score 33; DB 14; Length 126;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
DB 64 HGDLLLE 69

RESULT 14  
US-10-425-114-57360  
; Sequence 57360, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425.114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57360  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17312D02\_FLI pep  
; US-10-425-114-57360

Query Match 100.0%; Score 33; DB 12; Length 244;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
DB 231 HGDLLLE 236

RESULT 15  
US-10-353-690-88  
; Sequence 88, Application US/10353690  
; Publication No. US20030215840A1  
; GENERAL INFORMATION:  
; APPLICANT: Logan, Thomas Joseph  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Healy, Aileen  
; APPLICANT: Acton, Susan L.  
; APPLICANT: Donoghue, Mary

```

; APPLICANT: Staaliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodriguez-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: Cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 15489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1RNMIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-353-690-88

Query Match          100.0%; Score 33; DB 15; Length 379;
Best Local Similarity 100.0%; Pred.No.1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLE 6
        |||||
Db      93 HGDLE 98

Search completed: April 19, 2004, 12:54:58
Job time : 2.52632 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.3795 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pap.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description       |
|------------|-------|---------------|--------|----|-------------------|
| 1          | 47    | 100.0         | 584    | 1  | US-08-448-196A-7  |
| 2          | 47    | 100.0         | 585    | 1  | US-08-153-064-14  |
| 3          | 47    | 100.0         | 585    | 2  | US-08-448-196A-3  |
| 4          | 47    | 100.0         | 585    | 2  | US-08-984-176-1   |
| 5          | 47    | 100.0         | 585    | 2  | US-08-702-572-2   |
| 6          | 47    | 100.0         | 585    | 3  | US-08-769-746-2   |
| 7          | 47    | 100.0         | 585    | 4  | US-10-153-064-5   |
| 8          | 47    | 100.0         | 609    | 1  | US-08-222-619-3   |
| 9          | 47    | 100.0         | 609    | 1  | US-08-433-037-4   |
| 10         | 47    | 100.0         | 609    | 4  | US-08-897-956A-2  |
| 11         | 47    | 100.0         | 609    | 4  | US-10-153-064-7   |
| 12         | 47    | 100.0         | 609    | 4  | US-09-976-594-977 |
| 13         | 47    | 100.0         | 609    | 5  | PCT-US95-04075-3  |
| 14         | 47    | 100.0         | 610    | 2  | US-08-797-689-2   |
| 15         | 47    | 100.0         | 610    | 4  | US-09-384-186-2   |
| 16         | 47    | 100.0         | 651    | 4  | US-10-153-064-133 |
| 17         | 47    | 100.0         | 652    | 4  | US-10-153-064-96  |
| 18         | 47    | 100.0         | 652    | 4  | US-10-153-064-99  |
| 19         | 47    | 100.0         | 652    | 4  | US-10-153-064-105 |
| 20         | 47    | 100.0         | 652    | 4  | US-10-153-064-132 |
| 21         | 47    | 100.0         | 653    | 4  | US-10-153-064-131 |
| 22         | 47    | 100.0         | 656    | 4  | US-10-153-064-130 |
| 23         | 47    | 100.0         | 660    | 4  | US-10-153-064-90  |
| 24         | 47    | 100.0         | 660    | 4  | US-10-153-064-93  |
| 25         | 47    | 100.0         | 668    | 4  | US-10-153-064-102 |
| 26         | 47    | 100.0         | 676    | 4  | US-10-153-064-95  |
| 27         | 47    | 100.0         | 676    | 4  | US-10-153-064-98  |

Sequence 104, App  
Sequence 127, App  
Sequence 129, App  
Sequence 123, App  
Sequence 92, Appl  
Sequence 101, App  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 89, Appl  
Sequence 4, Appl  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 4, Appl  
US-10-153-064-104  
US-10-153-064-127  
US-10-153-064-129  
US-10-153-064-125  
US-10-153-064-123  
US-10-153-064-92  
US-10-153-064-101  
US-08-256-938-2  
US-08-256-938-4  
US-08-797-689-16  
US-09-984-186-16  
US-08-897-956A-3  
US-10-153-064-89  
US-08-448-196A-4  
US-08-448-196A-5  
US-08-448-196A-6  
US-08-134-638-1  
US-08-790-912-4

#### ALIGNMENTS

RESULT 1  
US-08-448-196A-7  
Sequence 7, Application US/08448196A  
Patent No. 5780594  
GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C.  
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
RELATED PROTEINS  
TITLE OF INVENTION: RELATED PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NASA  
STREET: MARSHALL SPACE FLIGHT CENTER  
CITY: HUNTSVILLE  
STATE: ALABAMA  
COUNTRY: USA  
ZIP: 35812  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,196A  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BROAD JR., ROBERT L.  
REGISTRATION NUMBER: 18,757  
REFERENCE/DOCKET NUMBER: XX/WFS-28402-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 205-544-0021  
TELEFAX: 205-544-0258  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 584 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-448-196A-7

Query Match 100.0%; Score 47; DB 1; Length 584;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AKQEPERNE 9

|||||||

Db 92 AKQEPERNE 100

RESULT 2

US-08-153-799-14

; Sequence 14, Application US/08153799

; Patent No. 576883

; GENERAL INFORMATION:

; APPLICANT: Ballance, David J

; APPLICANT: Goodey, Andrew R

; TITLE OF INVENTION: Polypeptides

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: R Hain Swope, BOC Health Care Inc

; STREET: 100 Mountain Avenue

; CITY: Murray Hill

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07974

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/153.799

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/847975

; FILING DATE: 06-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 8909916.2

; FILING DATE: 29-APR-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB90/00650

; FILING DATE: 26-APR-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/775952

; FILING DATE: 29-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Swope, R Hain

; REGISTRATION NUMBER: 24864

; REFERENCE/DOCKET NUMBER: 92H832

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 665 2400

; TELEFAX: (908) 771 6159

; TELEX: 219484

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 585 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: Region

; LOCATION: 369..419

; OTHER INFORMATION: /note= "Alternative C-termini of

; OTHER INFORMATION: HSA(1-n)"

; FEATURE:

; NAME/KEY: Region

; LOCATION: 1..585

; OTHER INFORMATION: /note= "Amino acid sequence of

; OTHER INFORMATION: natural HSA"

US-08-153-799-14

Query Match 100.0%; Score 47; DB 1; Length 585;

Best Local Similarity 100.0%; Fred. No. 0.38;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 AKQEPERNE 100

Db 92 AKQEPERNE 100

RESULT 3

US-08-448-196A-3

; Sequence 3, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TITLE OF INVENTION: RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448.196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: BROAD JR., ROBERT L.

; REGISTRATION NUMBER: 18,757

; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 205-544-0021

; TELEFAX: 205-544-0258

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 585 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 100.0%; Score 47; DB 1; Length 585;

Best Local Similarity 100.0%; Fred. No. 0.38;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 AKQEPERNE 100

Db 92 AKQEPERNE 100

RESULT 4

US-08-984-176-1

; Sequence 1, Application US/08984176

; Patent No. 5948609

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C

; APPLICANT: HO, JOSEPH X

; APPLICANT: RUKER, FLORIAN

; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT

; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

; FILE REFERENCE: 08/984.176

; CURRENT APPLICATION NUMBER: US/08/984.176

; CURRENT FILING DATE: 1997-12-03

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match      100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 5
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Blewas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match      100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 6
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
```

```
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match      100.0%; Score 47; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 7
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 47; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100
```

RESULT 8  
US-08-222-619-3  
; Sequence 3, Application US/08222619  
; Patent No. 5652352  
; GENERAL INFORMATION:  
; APPLICANT: Lichenstein, Henri  
; APPLICANT: Lyons, David  
; APPLICANT: Wurfel, Mark  
; APPLICANT: Wright, Samuel  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RR  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,619  
; FILING DATE:  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
DB 116 AKQEPERNE 124

RESULT 9  
US-08-433-037-4  
; Sequence 4, Application US/08433037  
; Patent No. 5707828  
; GENERAL INFORMATION:  
; APPLICANT: Sreekrishna, Kotikanyadan  
; APPLICANT: Barr, Kathryn A.  
; APPLICANT: Brierley, Russell A.  
; APPLICANT: Thill, Gregory P.  
; APPLICANT: Tschopp, Juerg F.  
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
; TITLE OF INVENTION: PICHIA PASTORIS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11530-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/433,037  
; FILING DATE: 03-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9108Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
DB 116 AKQEPERNE 124

RESULT 10  
US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; PRIOR FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 47; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
DB 116 AKQEPERNE 124

RESULT 11  
US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 12
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 13
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/ARRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 14
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-36,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match      100.0%; Score 47; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match      100.0%; Score 47; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 14
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-36,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match      100.0%; Score 47; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Mon Apr 19 13:27:30 2004

QY 1 AKOEPERNE 9  
 |||||  
 Db 116 AKOEPERNE 124

RESULT 15

US-09-984-186-2  
 ; Sequence 2, Application US/09984186  
 ; Patent No. 6686179

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard  
 Fournier, Alain  
 Guitton, Jean-Dominique  
 Jung, Gerard  
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Road, 3043  
 CITY: Collegeville  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: Macintosh  
 OPERATING SYSTEM: System 7.1  
 SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/984,186  
 FILING DATE: 29-Oct-2001  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689  
 FILING DATE: 31-JAN-1997  
 APPLICATION NUMBER: US 08/256,927  
 FILING DATE: 28-JUL-1994  
 APPLICATION NUMBER: FR 92/01064  
 FILING DATE: 31-JAN-1992  
 APPLICATION NUMBER: PCT/FR93/00085  
 FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.  
 REGISTRATION NUMBER: P-39,619  
 REFERENCE/DOCKET NUMBER: ST92006-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (610) 454-3839  
 TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 610 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-984-186-2

Query Match 100.0%; Score 47; DB 4; Length 610;  
 Best Local Similarity 100.0%; Pred No. 0.39;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKOEPERNE 9  
 |||||  
 Db 116 AKOEPERNE 124

Search completed: April 19, 2004, 12:05:19  
 Job time : 1.3795 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.853186 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_170\_176

Perfect score: 32

Sequence: 1 QAAADKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: Pirl.\*

2: Pirl.\*

3: Pirl.\*

4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 32    | 100.0       | 600    | A47391   | serum albumin prec |
| 2          | 32    | 100.0       | 605    | 1 ABPGS  | serum albumin prec |
| 3          | 32    | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 4          | 29    | 90.6        | 96     | AH3273   | hypothetical cytos |
| 5          | 29    | 90.6        | 128    | JC1273   | ribosomal protein  |
| 6          | 29    | 90.6        | 133    | T11792   | ribosomal protein  |
| 7          | 29    | 90.6        | 311    | AG3133   | 2-hydroxyacid dehy |
| 8          | 29    | 90.6        | 311    | E96154   | hypothetical prote |
| 9          | 29    | 90.6        | 315    | D95295   | probable D-isomer  |
| 10         | 29    | 90.6        | 387    | E70716   | probable succinyl- |
| 11         | 29    | 90.6        | 608    | 2 S57632 | serum albumin prec |
| 12         | 29    | 90.6        | 906    | AD3267   | protein translocas |
| 13         | 28    | 87.5        | 260    | C83362   | hypothetical prote |
| 14         | 28    | 87.5        | 293    | B87692   | carbonic anhydrase |
| 15         | 28    | 87.5        | 409    | A43256   | mobilization prote |
| 16         | 28    | 87.5        | 462    | A86717   | conserved hypothet |
| 17         | 28    | 87.5        | 523    | B38145   | invariant surface  |
| 18         | 28    | 87.5        | 570    | AD0223   | flagellar M-ring p |
| 19         | 28    | 87.5        | 608    | A46312   | gag polyprotein -  |
| 20         | 28    | 87.5        | 619    | A41971   | surface protein ps |
| 21         | 28    | 87.5        | 619    | A97887   | surface protein ps |
| 22         | 28    | 87.5        | 919    | T37062   | probable transcrip |
| 23         | 28    | 87.5        | 1061   | D69799   | cyclochrome P450 / |
| 24         | 28    | 87.5        | 1337   | T30291   | dextranase - Strep |
| 25         | 27    | 84.4        | 60     | H95354   | protein [imported  |
| 26         | 27    | 84.4        | 130    | 1 R35C8  | ribosomal protein  |
| 27         | 27    | 84.4        | 130    | 2 C91150 | 30S ribosomal subu |
| 28         | 27    | 84.4        | 130    | 2 G85995 | 30S ribosomal subu |
| 29         | 27    | 84.4        | 130    | 2 AG1007 | 30S ribosomal chai |

ALIGNMENTS

RESULT 1

A47391

serum albumin precursor - rhesus macaque

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999

C;Accession: A47391

R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwalet, J.; Putnam, F

Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993

A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin

A;Reference number: A47391; MUID:93211971; PMID:8460152

A;Contents: B/B homozygote

A;Accession: A47391

A;Status: preliminary

A;Molecule type: mRNA; protein

A;Residues: 1-600 <WAT>

A;Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PIR:G342295

A;Experimental source: liver

A;Note: sequence extracted from NCBI backbone (NCBI:128280, NCBIP:128281)

C;Superfamily: serum albumin; serum albumin repeat homology

F;21-194/Domain: serum albumin repeat homology <SA1>

F;213-386/Domain: serum albumin repeat homology <SA2>

F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 32; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAADKAA 7

Db 186 QAAADKAA 192

RESULT 2

ABPGS

serum albumin precursor - pig (fragment)

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999

C;Accession: S01382; A61006

R;Weinstock, J.; Baldwin, G. S.

Nucleic Acids Res. 16, 9045, 1988

A;Title: Nucleotide sequence of porcine liver albumin.

A;Reference number: S01382; MUID:89016582; PMID:3174440

A;Accession: S01382

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-605 <WEI>

A;Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798

R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.

J. Bone Miner. Res. 4, 235-241, 1989

A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of minera

A;Reference number: A61006; MUID:89269769; PMID:2728927

A;Accession: A61006

outer membrane pro  
outer membrane pro  
histone-like prote  
hypothetical prote  
outer membrane pro  
cationic 19 kDa ou  
MHC cell surface a  
hypothetical prote  
hypothetical prote  
sorbose-permease P  
probable sorbose P  
phosphotransferase  
phosphotransferase  
mannose-specific P  
PTS enzyme IIC, ma  
phosphotransferase

A:Molecule type: protein  
A:Residues: 23-51,'X',53-54,'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LIM>  
A:Experimental source: dental enamel  
A>Note: albumin and other serum proteins are also found in bone  
A:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membra  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-22/Domain: propeptide #status predicted <PRO>  
F:23-605/Product: serum albumin #status predicted <MAT>  
F:27-199/Domain: serum albumin repeat homology <SA1>  
F:218-391/Domain: serum albumin repeat homology <SA2>  
F:410-589/Domain: serum albumin repeat homology <SA3>  
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4  
F:266/Binding site: bilirubin (lys) #status predicted

Query Match 100.08; Score 32; DB 1; Length 605;  
Best Local Similarity 100.08; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAADKAA 7  
Db 191 QAAADKAA 197  
|||||

RESULT 3  
ASHUS  
serum albumin precursor [validated] - human  
N/Alternate names: preproalbumin  
N/Contains: kinetensin  
C/Species: Homo sapiens (man)  
C/Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000  
C/Accession: A93743; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S39  
R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli  
A:Reference number: A93743; MUID:82081982; PMID:6171778  
A:Accession: A93743  
A:Molecule type: mRNA  
A:Residues: 1-419,'K',421-609 <LAW>  
A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; GB:I00133; NID:G28591; PIDN:CAA23  
R/Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A:Reference number: A93936; MUID:82105994; PMID:6275391  
A:Accession: A93936  
A:Molecule type: mRNA  
A:Residues: 1-120,'G',122-609 <DUG>  
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
R/Drano, V.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A:Reference number: I39427; MUID:86140099; PMID:2419329  
A:Accession: I39427  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <URA>  
A:Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173  
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A:Reference number: I59286; MUID:94181575; PMID:8134387  
A:Accession: I59286  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 282-290,'KSPFDIQ',<WAT>  
A:Cross-references: GB:869192; NID:G546032; PIDN:AB30282.1; PID:G546033  
A>Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A:Reference number: I59313; MUID:94294404; PMID:8022807

A/Accession: I59313  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 589-590,'ALPRRVKNLLQVQLP',<MAD>  
A:Cross-references: GB:S70799; NID:G547231; PIDN:AB31177.1; PID:G547232  
A>Note: this frame-shift variant is designated albumin Bazzano; four additional variant:  
R/Nenaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: G08292  
A:Accession: G01747  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-120,'G',122-455 <MEN>  
A:Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex  
A:Reference number: S55314; MUID:95275251; PMID:7755581  
A:Accession: S55314  
A:Molecule type: protein  
A:Residues: 19-27 <LED>  
R/Meloun, B.; Moravek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A:Title: Complete amino acid sequence of human serum albumin.  
A:Reference number: A91420; MUID:76187907; PMID:1225573  
A:Accession: A91420  
A:Molecule type: protein  
A:Residues: 23-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395  
R/Roehr, U.; Spittler, G.; Irprier, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A:Reference number: S06422  
A>Note: this paper is in German, with an English abstract  
A:Accession: S06422  
A:Molecule type: protein  
A:Residues: 25-48 <ROE>  
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A:Title: Mass spectrometric identification of modifications to human serum albumin treat  
A:Reference number: S36882; MUID:93384321; PMID:8373198  
A:Accession: S36882  
A:Molecule type: protein  
A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>  
R/Kausler, E.; Spittler, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A:Title: Bruchstücke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelm  
A:Reference number: S17599; MUID:92126241; PMID:1772598  
A:Accession: S17599  
A:Molecule type: protein  
A:Residues: 25-54;354-357;431-447 <KAU>  
A>Note: 49-Leu was also found  
R/Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid protea  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: A45800  
A:Molecule type: protein  
A:Residues: 166-173,'L',<MOG>  
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tr  
A:Reference number: A03239; MUID:86242180; PMID:3087352  
A:Accession: A03239  
A:Molecule type: protein  
A:Residues: 166-173,'L',<MOG>  
R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins,  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A:Title: Mutations in genetic variants of human serum albumin found in Italy.  
A:Reference number: A38255; MUID:91062352; PMID:2247440  
A:Accession: C38255  
A:Molecule type: protein  
A:Residues: 76-111 <GAL1>  
A:Accession: B38255

A:Molecule type: protein  
A:Residues: 82-105,'K',107-110 <GAL2>  
A>Note: this variant is designated albumin Vibo Valentia  
A:Accession: A38255  
A:Molecule type: protein  
A:Residues: 76-83,'K',85-106 <GAL3>  
A>Note: this variant is designated albumin Torino  
R:Minichioti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A:Title: The structural characterization and bilirubin-binding properties of albumin Her  
A:Reference number: S33298; MUID:93292504; PMID:8513793  
A:Accession: S33298  
A:Molecule type: protein  
A:Residues: 255-263,'E',265-281 <MIN1>  
A>Note: this variant is designated albumin Herborn  
R:Minichioti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A:Title: Two albumins with identical electrophoretic mobility are produced by differ  
A:Reference number: S21078; MUID:92190239; PMID:1347703  
A:Accession: S21078  
A:Molecule type: protein  
A:Residues: 354-356,'K',358-378 <MIN2>  
A>Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,  
R:He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A:Title: Atomic structure and chemistry of human serum albumin.  
A:Reference number: A46756; MUID:92334427; PMID:1630489  
A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
R:Brown, J.R.; Shockley, P.; Behrens, P.Q.  
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
A:Reference number: A94442  
A:Contents: annotation; three-dimensional structure and disulfide bonds  
R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A:Title: Disulfide bonds in human serum albumin.  
A:Reference number: A30930  
A:Contents: annotation; disulfide bonds  
R:Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A:Reference number: A90299; MUID:78186630; PMID:656055  
A:Contents: annotation; bilirubin-binding site  
R:Peters, T.; Reed, R.G.  
In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20,  
A:Title: Serum albumin: conformation and active sites.  
A:Reference number: A94408  
A:Contents: annotation; binding sites  
R:Harper, M.E.; Dugaiczky, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
A:Reference number: A90028; MUID:83279982; PMID:6192711  
A:Contents: annotation; gene position  
R:Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
A:Reference number: A46755; MUID:76257808; PMID:955075  
A:Contents: annotation  
A>Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid  
R:Bohney, J.P.; Fonda, M.L.; Feidhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat  
A:Reference number: A56294; MUID:92183881; PMID:1544460  
A:Contents: annotation  
A>Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p  
ase activity  
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
C:Comment: A large number of variants of human serum albumin have been described.  
C:Genetics:  
A:Gene: GDB:ALB  
A:Cross-references: GDB:118990; OMIM:103600  
A:Map position: 4q11-4q13  
C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridoxal  
F:19-Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-609/Product: serum albumin #status experimental <MPT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:166-174/Product: kinetensin #status experimental <KIP>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:421-592/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental  
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 32; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
|||  
DB 194 QAADKAA 200

RESULT 4  
AH3273  
Hypothetical cytosolic protein BMS10173 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AH3273  
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes;  
Proc.Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten  
A:Reference number: AD3252; PMID:11756868  
A:Accession: AH3273  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51355.1; PID:gi7982054; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BME10173  
A:Map position: 1

Query Match 90.6%; Score 29; DB 2; Length 96;  
Best Local Similarity 85.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
|||  
DB 56 EAADKAA 62

RESULT 5  
JC1273  
ribosomal protein L7/L12 - Streptomyces antibioticus  
C:Species: Streptomyces antibioticus  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
C:Accession: JC1273  
R:Parra, F.; Blanco, G.; Alonso, J.M.M.; Balbin, M.; Mendez, C.; Salas, J.A.  
Gene 118, 127-129, 1992  
A:Title: Cloning and sequence of a gene encoding the L7/L12 ribosomal protein equivalent  
A:Reference number: JC1273; MUID:92380478; PMID:1511874  
A:Accession: JC1273  
A:Molecule type: DNA  
A:Residues: 1-128 <PAR>  
A:Cross-references: GB:M89911; NID:gi53436; PIDN:AAA26811.1; PID:gi53438  
C:Superfamily: Escherichia coli ribosomal protein L12  
C:Keywords: protein biosynthesis; ribosome

Query Match 90.6%; Score 29; DB 2; Length 128;  
Best Local Similarity 85.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7

Db 109 EAADKAA 115  
:|||||  
RESULT 6  
ribosomal protein L12 - Streptomyces virginiae  
T11792  
C:Species: Streptomyces virginiae  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T11792  
R:Yamada, Y.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: Z17345  
A:Accession: T11792  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-133 <YAM>  
A:Cross-references: EMBL:D50624  
C:Genetics:  
A:Gene: rplL  
C:Superfamily: Escherichia coli ribosomal protein L12  
C:Keywords: ribosome  
Query Match 90.6%; Score 29; DB 2; Length 133;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QAADKAA 7  
:|||||  
Db 114 EAADKAA 120  
:|||||  
RESULT 7  
2-hydroxyacid dehydrogenase Atu4691 [imported] - Agrobacterium tumefaciens (strain C58,  
AG3133  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AG3133  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AG3133  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-311 <KUR>  
A:Cross-references: GB:AE008689; PIDN:AAL45485.1; PID:gl7743192; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu4691  
A:Map position: linear chromosome  
C:Superfamily: phosphoglycerate dehydrogenase  
Query Match 90.6%; Score 29; DB 2; Length 311;  
Best Local Similarity 85.7%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QAADKAA 7  
:|||||  
Db 30 EAADKAA 36  
:|||||  
RESULT 8  
E98154  
hypothetical protein AGR\_L379 [imported] - Agrobacterium tumefaciens (strain C58, Cerec  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
C:Accession: E98154  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: E98154  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-311 <KUR>  
A:Cross-references: GB:AE007870; PIDN:AAK88759.1; PID:gl5158503; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L\_379  
A:Map position: linear chromosome  
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 90.6%; Score 29; DB 2; Length 311;  
Best Local Similarity 85.7%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
:|||||  
Db 30 EAADKAA 36  
:|||||

RESULT 9  
D95295  
probable D-isomer specific 2-hydroxyacid [imported] - Sinorhizobium meliloti (strain 10  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: D95295  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: D95295  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-315 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AAK64926.1; PID:gl4523348; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMA0510  
A:Genome: plasmid  
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 90.6%; Score 29; DB 2; Length 315;  
Best Local Similarity 85.7%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
:|||||  
Db 31 EAADKAA 37  
:|||||

RESULT 10  
E70716  
probable succinyl-coa synthetase - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: E70716  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,  
; Connor, R.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, S.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: E70716  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-387 <COL>  
A;Cross-references: GB:Z79700; GB:AL123456; NID:g3261628; PIDN:CAB01999.1; PID:g1524211  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: succ  
C;Superfamily: succinate-CoA ligase (ADP-forming) beta chain

Query Match 90.6%; Score 29; DB 2; Length 387;  
Best Local Similarity 85.7%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
:|||||

DB 376 EAADKAA 382  
:|||||

RESULT 11

serum albumin precursor - cat

C;Species: Felis silvestris catus (domestic cat)

C;Date: 18-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999

C;Accession: J04660; S57632

R;Hilger, C.; Grigioni, F.; Hentges, F.

Gene 169, 295-296, 1996

A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.

A;Reference number: J04660; MUID:96194824; PMID:8647469

A;Accession: J04660

A;Molecule type: mRNA

A;Residues: 1-608 <H12>

A;Cross-references: EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g886485

A;Experimental source: liver

C;Comment: This protein is the major protein component in plasma. It functions as a multi-

ein has 35 conserved cysteine residues.

C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: liver; plasma

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-24/Domain: propeptide #status predicted <PRP>

F;25-608/Product: serum albumin #status predicted <MAT>

F;29-202/Domain: serum albumin repeat homology <SA1>

F;221-394/Domain: serum albumin repeat homology <SA2>

F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 90.6%; Score 29; DB 2; Length 608;  
Best Local Similarity 85.7%; Pred. No. 16+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
:|||||

DB 194 EAADKAA 200  
:|||||

RESULT 12

protein translocase, chain secA [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002

C;Accession: AD3267

R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AD3267

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-905 <KUR>

A;Cross-references: GB:AE008917; PIDN:AAL51303.1; PID:g17981998; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME10121

A;Map position: 1

C;Superfamily: preproteins translocase secA

Query Match 90.6%; Score 29; DB 2; Length 906;  
Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
:|||||

DB 736 EAADKAA 742  
:|||||

RESULT 13

C83362

hypothetical protein PA2260 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C;Accession: C83362

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bi-

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path-

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83362

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-260 <STO>

A;Cross-references: GB:AE004652; GB:AE004091; NID:g9948287; PIDN:AAG05648.1; GSPDB:GN00:

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2260

Query Match 87.5%; Score 28; DB 2; Length 260;  
Best Local Similarity 85.7%; Pred. No. 76;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
:|||||

DB 170 QAADKAA 176  
:|||||

RESULT 14

carbonic anhydrase [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: B87692

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo-

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M

Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: B87692

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-293 <STO>

A;Cross-references: GB:AE005673; NID:g13425312; PIDN:AAK25534.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC3572

Query Match 87.5%; Score 28; DB 2; Length 293;  
Best Local Similarity 85.7%; Pred. No. 86;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
:|||||

DB 64 QAADKAA 70  
:|||||

RESULT 15  
A43256  
mobilization protein mobA - Thiobacillus ferrooxidans plasmid pTF-PC2  
C:Species: Thiobacillus ferrooxidans  
C:Date: 10-Jun-1993 #sequence\_revision 01-Dec-1995 #text\_change 08-Oct-1999  
C:Accession: A43256; S27622  
R:Rohrer, J.; Rawlings, D.E.  
J. Bacteriol. 174, 6230-6237, 1992  
A:Title: Sequence analysis and characterization of the mobilization region of a broad-host-range plasmid from Thiobacillus ferrooxidans  
A:Reference number: A43256; MUID:93015664; PMID:1400173  
A:Accession: A43256  
A:Status: Preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-409 <ROH>  
A:Cross-references: EMBL:MS7717; NID:G154659; PIDN:AAA27389.1; PID:G154660  
A:Note: sequence extracted from NCBI backbone (NCBIP:115305)  
C:Genetics:  
A:Gene: mobA  
A:Genome: plasmid

Query Match 87.5%; Score 28; DB 2; Length 409;  
Best Local Similarity 85.7%; Pred.No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
| | | | |  
DB 306 QAEXKAA 312

Search completed: April 19, 2004, 12:02:25  
Job time : 1.85319 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 5.29363 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_280\_288

Perfect score: 47

Sequence: 1 EXPLEKSH 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 47    | 100.0       | 112    | 4     | AA02636 Human pol  |
| 2          | 47    | 100.0       | 114    | 4     | AA025284 Human pro |
| 3          | 47    | 100.0       | 123    | 4     | AA023861 abcEST en |
| 4          | 47    | 100.0       | 126    | 4     | AA004435 Human pol |
| 5          | 47    | 100.0       | 188    | 3     | AA083948 Yeast cod |
| 6          | 47    | 100.0       | 293    | 4     | AAU33082 Novel hum |
| 7          | 47    | 100.0       | 303    | 2     | AA014178 Human ser |
| 8          | 47    | 100.0       | 373    | 1     | AA090387 N-termina |
| 9          | 47    | 100.0       | 388    | 1     | AA090389 N-termina |
| 10         | 47    | 100.0       | 389    | 1     | AA090390 N-termina |
| 11         | 47    | 100.0       | 390    | 1     | AA090391 N-termina |
| 12         | 47    | 100.0       | 407    | 1     | AA090392 N-termina |
| 13         | 47    | 100.0       | 463    | 2     | AA014179 Human ser |
| 14         | 47    | 100.0       | 500    | 7     | ADD32019 Heterolog |
| 15         | 47    | 100.0       | 550    | 4     | AA029877 Novel hum |
| 16         | 47    | 100.0       | 584    | 6     | AB072381 Mature hu |
| 17         | 47    | 100.0       | 585    | 1     | AA093344 Sequence  |
| 18         | 47    | 100.0       | 585    | 1     | AA090388 Mature hu |
| 19         | 47    | 100.0       | 585    | 1     | AA091422 Human nor |
| 20         | 47    | 100.0       | 585    | 2     | AA05318 Human ser  |
| 21         | 47    | 100.0       | 585    | 2     | AA080518 Human ser |
| 22         | 47    | 100.0       | 585    | 2     | AA026207 Human ser |
| 23         | 47    | 100.0       | 585    | 2     | AA026362 Synthetic |
| 24         | 47    | 100.0       | 585    | 2     | AA020029 Human ser |
| 25         | 47    | 100.0       | 585    | 2     | AA080301 Human ser |

|    |    |       |     |   |                    |
|----|----|-------|-----|---|--------------------|
| 26 | 47 | 100.0 | 585 | 2 | AA020111 HSA prote |
| 27 | 47 | 100.0 | 585 | 2 | AA059841 Mature pr |
| 28 | 47 | 100.0 | 585 | 3 | AA084873 Amino aci |
| 29 | 47 | 100.0 | 585 | 3 | AA083946 Yeast cod |
| 30 | 47 | 100.0 | 585 | 4 | AA052567 Mature hu |
| 31 | 47 | 100.0 | 585 | 4 | AA012417 Human alb |
| 32 | 47 | 100.0 | 585 | 4 | AA012403 Human alb |
| 33 | 47 | 100.0 | 585 | 4 | AA013129 Human alb |
| 34 | 47 | 100.0 | 585 | 4 | AA013135 Human alb |
| 35 | 47 | 100.0 | 585 | 4 | AA013311 Human alb |
| 36 | 47 | 100.0 | 585 | 4 | AA013399 Human alb |
| 37 | 47 | 100.0 | 585 | 4 | AA079006 Human mat |
| 38 | 47 | 100.0 | 585 | 4 | AA08578 Human ser  |
| 39 | 47 | 100.0 | 585 | 5 | AAU75220 Mature fo |
| 40 | 47 | 100.0 | 585 | 5 | ABJ00986 B lymphoc |
| 41 | 47 | 100.0 | 585 | 5 | ABG63321 Human ser |
| 42 | 47 | 100.0 | 585 | 5 | ABG33847 Human B L |
| 43 | 47 | 100.0 | 585 | 5 | ABG71291 Glycosyla |
| 44 | 47 | 100.0 | 585 | 6 | ABR55695 Human alb |
| 45 | 47 | 100.0 | 585 | 6 | AA030936 Human alb |

ALIGNMENTS

RESULT 1

AA002636

ID AA002636 standard; protein; 112 AA.

XX AC AA002636;

XX XX 06-NOV-2001 (first entry)

XX XX Human polypeptide SEQ ID NO 16528.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX XX 07-SEP-2001.

XX XX 26-FEB-2001; 2001WO-US004927.

XX XX 28-FEB-2000; 2000US-00515125.

XX XX 18-MAY-2000; 2000US-00577409.

XX XX (HYSE-) HYSEQ INC.

XX XX Tang YT, Liu C, Drmanac RT;

XX XX WPI; 2001-514838/56.

XX XX N-PSDB; AA02557.

XX XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
and treating e.g. leukemia, inflammation and immune disorders.

XX XX Claim 20; SEQ ID NO 16528; 1399pp + Sequence Listing; English.

XX XX The invention relates to human polynucleotides (AA079941-AA093841) and  
the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
cytokine, cell proliferation or cell differentiation or which may induce  
production of other cytokines in other cell populations. The  
polynucleotides and polypeptides are useful in gene therapy, vaccines or  
peptide therapy. The polypeptides have various cytokine-like activities,  
e.g. stem cell growth factor activity, haematopoiesis regulating  
activity, tissue growth factor activity, immunomodulatory activity and  
activin/inhibin activity and may be useful in the diagnosis and/or  
treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 112 AA;

Query Match 100.0%; Score 47; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKPLLEKSH 9  
 DB 3 EKPLLEKSH 11  
 |||||

RESULT 2  
 AAM25284  
 ID AAM25284 standard; protein; 114 AA.  
 XX  
 AC AAM25284;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:799.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 antibacterial; endocrine; cardiant; central nervous system; virucide;  
 anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
 antiaggregant; haemostatic; vulnary; antitumor; osteopathic; eczema;  
 dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
 neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;  
 antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
 allergic rhinitis; diabetes; multiple sclerosis; depression;  
 Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 neurological disorder.

XX Homo sapiens.  
 OS  
 XX  
 PN WO200153455-A2.  
 XX  
 XX  
 PD 26-JUL-2001.  
 XX  
 XX  
 PF 22-DEC-2000; 2000WO-US035017.  
 XX  
 PF 23-DEC-1999; 99US-004712175.  
 PR  
 PR 21-JAN-2000; 2000US-00488725.  
 PR  
 PR 25-APR-2000; 2000US-00552317.  
 XX  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 XX  
 DR WPI; 2001-457603/49.  
 DR N-PSDB; AAH99225.  
 XX  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
 XX  
 XX  
 PS Claim 20; Page 186; 1217pp; English.

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and cells  
 CC they are expressed in, such as: antiinflammatory; antirheumatic;  
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;  
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnary;  
 CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antitense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders  
 XX  
 SQ Sequence 114 AA;

Query Match 100.0%; Score 47; DB 4; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKPLLEKSH 9  
 DB 79 EKPLLEKSH 87  
 |||||

RESULT 3  
 AAM23861  
 ID AAM23861 standard; protein; 123 AA.  
 XX  
 AC AAM23861;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE abCEST encoded protein SEQ ID NO: 1386.

Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;  
 KW gene therapy; nutrition.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200154477-A2.  
 XX  
 XX  
 PD 02-AUG-2001.  
 XX  
 XX  
 PF 25-JAN-2001; 2001WO-US002687.  
 XX  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 PR  
 PR 17-JUL-2000; 2000US-00617746.  
 PR  
 PR 03-AUG-2000; 2000US-00631451.  
 PR  
 PR 15-SEP-2000; 2000US-00683870.  
 XX  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX  
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 XX  
 XX  
 DR WPI; 2001-476164/51.  
 DR N-PSDB; AAH98520.  
 XX  
 XX  
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use.  
 XX  
 PS Claim 20; Page 975; 1275pp; English.

The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a



CC protein of the invention  
XX  
SQ Sequence 123 AA;  
  
Query Match 100.0%; Score 47; DB 4; Length 123;  
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 EKPLLEKSH 9  
Db 77 EKPLLEKSH 85  
  
RESULT 4  
AAO04435  
ID AAO04435 standard; protein; 126 AA.  
XX  
AC AAO04435;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 18327.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US004927.  
XX  
PR 28-FEB-2000; 2000US-00515126.  
XX  
PR 18-MAY-2000; 2000US-00577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
XX  
DR N-PSDB; AAI84366.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
PS Claim 20; SEQ ID NO 18327; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pot\_sequences  
XX  
SQ Sequence 126 AA;  
  
Query Match 100.0%; Score 47; DB 4; Length 126;  
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 EKPLLEKSH 9  
Db 77 EKPLLEKSH 85

Db 118 EKPLLEKSH 126  
  
RESULT 5  
AAV83948  
ID AAV83948 standard; protein; 188 AA.  
XX  
AC AAV83948;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Yeast codon-biased recombinant HSA protein fragment HSA-II.  
XX  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX  
OS Homo sapiens.  
XX  
OS Synthetic.  
XX  
PN CNI239103-A.  
XX  
PD 22-DEC-1999.  
XX  
PF 17-JUN-1998; 98CN-00102506.  
XX  
PR 17-JUN-1998; 98CN-00102506.  
XX  
PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
XX  
PI Li S, Lu D;  
XX  
DR WPI; 2000-351198/31.  
XX  
DR N-PSDB; AAA10093.  
XX  
PT Process for preparing recombinant human serum albumin comprising yeast  
PT biased sex codons - uses a recombinant DNA technique.  
XX  
PS Example 1; Fig 5; 44pp; Chinese.  
XX  
CC The method relates to a method of recombinantly producing human serum  
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
CC a yeast codon bias. The complete HSA gene (AA10091) was generated as  
CC three synthetic fragments (AAA10092-AA10094) joined by recombinant DNA  
CC technology. Each HSA fragment was synthesised from overlapping  
CC oligonucleotide fragments that were extended. This sequence represents  
CC the sequence of the HSA fragment HSA-II encoded by the human gene with a  
CC yeast codon bias. The invention also covers a recombinant expression  
CC vector, yeast host cells carrying the recombinant expression vector and  
CC the process for producing human serum albumin in the yeast host cell,  
CC especially in secretory mode  
XX  
SQ Sequence 188 AA;  
  
Query Match 100.0%; Score 47; DB 3; Length 188;  
Best Local Similarity 100.0%; Pred. No. 0.19; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 EKPLLEKSH 9  
Db 96 EKPLLEKSH 104  
  
RESULT 6  
AAU33082  
ID AAU33082 standard; protein; 293 AA.  
XX  
AC AAU33082;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3573.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX Homo sapiens.  
XX WO200179449-A2.  
XX 25-OCT-2001.  
XX 16-APR-2001; 2001WO-US008656.  
XX 18-APR-2000; 2000US-00552929.  
XX 26-JAN-2001; 2001US-00770160.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.  
XX Claim 20; Page 705; 765pp; English.  
XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU33304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX Sequence 293 AA;  
Query Match 100.0%; Score 47; DB 4; Length 293;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EKPLLEKSH 9  
DB 134 EKPLLEKSH 142  
RESULT 7  
ID AAR14178 standard; protein; 303 AA.  
XX AAR14178;  
XX AAR14178;  
XX 19-DEC-1991 (first entry)  
XX Human serum albumin lacking C-terminal fragment.  
XX HSA; fusion protein; drug.  
XX Homo sapiens.  
XX JP03201987-A.  
XX 03-SEP-1991.  
XX 29-DEC-1989; 89JP-00344701.  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX Homo sapiens.  
XX WO200179449-A2.  
XX 25-OCT-2001.  
XX 16-APR-2001; 2001WO-US008656.  
XX 18-APR-2000; 2000US-00552929.  
XX 26-JAN-2001; 2001US-00770160.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.  
XX Claim 20; Page 705; 765pp; English.  
XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU33304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX Sequence 293 AA;  
Query Match 100.0%; Score 47; DB 4; Length 293;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EKPLLEKSH 9  
DB 134 EKPLLEKSH 142  
RESULT 7  
ID AAR14178 standard; protein; 303 AA.  
XX AAR14178;  
XX AAR14178;  
XX 19-DEC-1991 (first entry)  
XX Human serum albumin lacking C-terminal fragment.  
XX HSA; fusion protein; drug.  
XX Homo sapiens.  
XX JP03201987-A.  
XX 03-SEP-1991.  
XX 29-DEC-1989; 89JP-00344701.

XX 29-DEC-1989; 89JP-00344701.  
XX (TOFU) TONEN CORP.  
XX WPI; 1991-300976/41.  
XX Human serum albumin fragment - where C-terminal of human serum albumin is  
XX lacking and which can be combined with various drugs.  
XX Claim 1; Page 1; 23pp; Japanese.  
XX This sequence corresponds to amino acids 1 to 303 of mature human serum  
XX albumin. The fragment lacking the C-terminal sequence can form part of a  
XX fusion protein, for example with drugs. (This sequence is taken from the  
XX full-length HSA sequence in EP-330451). See also AAR14179  
XX Sequence 303 AA;  
Query Match 100.0%; Score 47; DB 2; Length 303;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EKPLLEKSH 9  
DB 280 EKPLLEKSH 288  
RESULT 8  
ID AAP90387 standard; protein; 373 AA.  
XX AAP90387;  
XX 24-OCT-2003 (revised)  
XX 25-MAR-2003 (revised)  
XX 01-NOV-1989 (first entry)  
XX N-terminal of human serum albumin polypeptide.  
XX Human serum albumin polypeptide; plasma expanders.  
XX Homo sapiens; (Human).  
XX EP322094-A.  
XX 28-JUN-1989.  
XX 25-OCT-1988; 88EP-00310000.  
XX 30-OCT-1987; 87GB-00025529.  
XX (DELZ) DELTA BIOTECHNOLOGY LTD.  
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
XX WPI; 1989-186464/26.  
XX New N-terminal fragments of human serum albumin - esp. useful as blood  
XX plasma expanders.  
XX Claim 2; Page 9; 20pp; English.  
XX N-terminal portion of human serum albumin. Used as plasma expanders, or  
XX as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-  
XX MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS  
XX field)  
XX Sequence 373 AA;  
Query Match 100.0%; Score 47; DB 1; Length 373;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 47; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 0.43; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288

RESULT 12
AAP90392
ID AAP90392 standard; protein; 407 AA.
XX AC AAP90392;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 01-NOV-1989 (first entry)
XX DE N-terminal human serum albumin.
XX KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX OS Homo sapiens; (Human).
XX FN EP322094-A.
XX PN 28-JUN-1989.
XX PD 25-OCT-1988; 88EP-00310000.
XX PF 30-OCT-1987; 87GB-00025529.
XX PR (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX PA Ballance DJ, Hinchliffe E, Geisow MJ, Senior PU;
XX PI WPI; 1989-186464/26.
XX PS New N-terminal fragments of human serum albumin - esp. useful as blood
XX PT plasma expanders.
XX PS Claim 2; Page 9; 20pp; English.
XX CC N-terminal portion of human serum albumin. Used to make plasma expanders,
XX CC or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
XX CC -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX CC field)
XX SQ Sequence 407 AA;
Query Match 100.0%; Score 47; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 0.45; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288

RESULT 13
AAR14179
ID AAR14179 standard; protein; 463 AA.
XX AC AAR14179;
XX DT 19-DEC-1991 (first entry)
XX DE Human serum albumin lacking N-terminal fragment.

XX KW HSA; fusion protein; drug.
XX OS Homo sapiens.
XX PN JP03201987-A.
XX PD 03-SEP-1991.
XX PF 29-DEC-1989; 89JP-00344701.
XX PR 29-DEC-1989; 89JP-00344701.
XX PA (TOFU ) TONEN CORP.
XX DR WPI; 1991-300976/41.
XX PT Human serum albumin fragment - where C-terminal of human serum albumin is
XX PS lacking and which can be combined with various drugs.
XX PS Claim 6; Page 1; 23pp; Japanese.
XX CC This sequence corresponds to amino acids 123 to 585 of mature human serum
XX CC albumin. The fragment lacking the N-terminal sequence can form part of a
XX CC fusion protein, for example with drugs. (This sequence is taken from the
XX CC full-length HSA sequence in EP-330451). See also AAR14178
XX SQ Sequence 463 AA;
Query Match 100.0%; Score 47; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 0.51; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 158 EKPLLEKSH 166

RESULT 14
ADD32019
ID ADD32019 standard; protein; 500 AA.
XX AC ADD32019;
XX DT 15-JAN-2004 (first entry)
XX DE Heterologous fusion protein related protein sequence SEQ ID NO:25.
XX KW heterologous fusion protein;
XX KW hyperglycosylated granulocyte-colony stimulating factor; G-CSF;
XX KW human albumin; human albumin analogue; immunoglobulin; Fc;
XX KW immunostimulant; protein therapy; neutrophil level;
XX KW insufficient circulating neutrophil level;
XX KW chronic congenital neutropenia.
XX OS Synthetic.
XX PN WO2003076567-A2.
XX PD 18-SEP-2003.
XX PF 21-FEB-2003; 2003WO-US003120.
XX PR 05-MAR-2002; 2002US-0361948P.
XX PA (ELIL ) LILLY & CO ELI.
XX PI Beals JM, Kuchibhotla U;
XX DR WPI; 2003-902770/82.
XX PT New heterologous fusion proteins with granulocyte-colony stimulating
factor activity, useful for increasing neutrophil levels and treating

```

PT patients with low circulating neutrophils, such as after chemotherapy or  
 PT in neutropenia.  
 XX  
 PS Disclosure; SEQ ID NO 25; 126pp; English.  
 XX  
 CC The present invention describes a heterologous fusion protein (I)  
 CC comprising a hyperglycosylated granulocyte-colony stimulating factor (G-  
 CC CSF) analogue fused to a polypeptide having human albumin, human albumin  
 CC analogues, or fragments of human albumin, or the Fc portion of an  
 CC immunoglobulin, an analogue of the Fc portion of an immunoglobulin, or  
 CC fragments of the Fc portion of an immunoglobulin. Also described: (1) a  
 CC heterologous fusion protein, which is the product of the expression in a  
 CC host cell of an exogenous DNA sequence encoding (1); (2) an isolated  
 CC nucleic acid sequence comprising: (a) polynucleotides encoding (1); or  
 CC (b) a polynucleotide, which comprises any of 15 DNA sequences each  
 CC comprising 1044 base pairs (see ADD31996 to ADD32010), fused to the DNA  
 CC encoding a protein (i.e. human albumin, human albumin analogue or  
 CC fragments of human albumin); (3) increasing neutrophil levels in a mammal  
 CC comprising the administration of (1); (4) pharmaceutical formulations  
 CC adapted for the treatment of patients with insufficient neutrophil levels  
 CC comprising any of (1); (5) a vector comprising the polynucleotide of (2);  
 CC (6) host cells comprising the vector of (5), or expressing at least one  
 CC heterologous protein; and (7) producing (1). (1) has immunostimulant  
 CC activity, and can be used in protein therapy. (1) can be used for  
 CC increasing neutrophil levels in a mammal. (1) are particularly useful in  
 CC medicaments for the treatment of patients with insufficient circulating  
 CC neutrophil levels, or for the manufacture of a medicament for the  
 CC treatment of patients with insufficient circulating neutrophil levels.  
 CC The fusion protein is particularly useful for treating conditions  
 CC treatable by stimulation of circulating neutrophils, such as after  
 CC chemotherapy regimens or in chronic congenital neutropenia. The present  
 CC sequence is used in the exemplification of the present invention.  
 XX

Sequence 500 AA;

Query Match 100.0%; Score 47; DB 7; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 0.56;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9  
 |||||  
 Db 469 EKPLEKSH 477

RESULT 15  
 AAU29877  
 ID AAU29877 standard; protein; 550 AA.

AAU29877;

18-DEC-2001 (first entry)

Novel human secreted protein #368.

Human; vaccination; gene therapy; nutritional supplement;  
 stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 immune suppression; immune stimulation; anti-inflammatory; leukaemia.

Hom sapiens.

WO200179449-A2.

25-OCT-2001.

16-APR-2001; 2001WO-US008656.

18-APR-2000; 2000US-00552929.

26-JAN-2001; 2001US-00770160.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

PI  
 XX

DR WPI; 2001-6111725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.

XX  
 PS Claim 20; Page 206; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention

XX Sequence 550 AA;

Query Match 100.0%; Score 47; DB 4; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9

|||||

Db 245 EKPLEKSH 253

Search completed: April 19, 2004, 11:51:20

Job time : 5.29363 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_280\_288

Perfect score: 47  
Sequence: 1 EKPILLEKSH 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_5/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_5/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_5/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_5/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_5/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_5/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_5/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_5/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_5/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_5/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_5/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_5/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_5/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_5/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_5/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_5/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_5/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_5/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 47    | 100.0       | 114    | 12    | US-10-296-115-799  |
| 2          | 47    | 100.0       | 585    | 9     | Sequence 799, Appl |
| 3          | 47    | 100.0       | 585    | 10    | Sequence 2, Appl   |
| 4          | 47    | 100.0       | 585    | 10    | Sequence 445, Appl |
| 5          | 47    | 100.0       | 585    | 10    | Sequence 26, Appl  |
| 6          | 47    | 100.0       | 585    | 10    | Sequence 18, Appl  |
| 7          | 47    | 100.0       | 585    | 10    | Sequence 18, Appl  |
| 8          | 47    | 100.0       | 585    | 10    | Sequence 445, Appl |
| 9          | 47    | 100.0       | 585    | 10    | Sequence 18, Appl  |
| 10         | 47    | 100.0       | 585    | 11    | Sequence 18, Appl  |
| 11         | 47    | 100.0       | 585    | 11    | Sequence 18, Appl  |
| 12         | 47    | 100.0       | 585    | 12    | Sequence 11, Appl  |
| 13         | 47    | 100.0       | 585    | 12    | Sequence 31, Appl  |
| 14         | 47    | 100.0       | 585    | 13    | Sequence 34, Appl  |
| 15         | 47    | 100.0       | 585    | 14    | Sequence 5, Appl   |

|    |    |       |     |    |                   |                    |
|----|----|-------|-----|----|-------------------|--------------------|
| 16 | 47 | 100.0 | 585 | 14 | US-10-319-263-1   | Sequence 1, Appl   |
| 17 | 47 | 100.0 | 585 | 14 | US-10-319-263-2   | Sequence 2, Appl   |
| 18 | 47 | 100.0 | 585 | 14 | US-10-414-469-1   | Sequence 1, Appl   |
| 19 | 47 | 100.0 | 585 | 14 | US-10-414-469-2   | Sequence 2, Appl   |
| 20 | 47 | 100.0 | 585 | 14 | US-10-413-831-1   | Sequence 1, Appl   |
| 21 | 47 | 100.0 | 585 | 14 | US-10-413-831-2   | Sequence 2, Appl   |
| 22 | 47 | 100.0 | 585 | 15 | US-10-413-832-1   | Sequence 1, Appl   |
| 23 | 47 | 100.0 | 585 | 15 | US-10-413-832-2   | Sequence 2, Appl   |
| 24 | 47 | 100.0 | 585 | 15 | US-10-414-386-1   | Sequence 1, Appl   |
| 25 | 47 | 100.0 | 585 | 15 | US-10-414-386-2   | Sequence 2, Appl   |
| 26 | 47 | 100.0 | 585 | 15 | US-10-233-675A-11 | Sequence 11, Appl  |
| 27 | 47 | 100.0 | 585 | 15 | US-10-462-262-26  | Sequence 26, Appl  |
| 28 | 47 | 100.0 | 604 | 10 | US-09-984-010-7   | Sequence 7, Appl   |
| 29 | 47 | 100.0 | 609 | 10 | US-09-919-039-370 | Sequence 370, Appl |
| 30 | 47 | 100.0 | 609 | 12 | US-10-609-346-12  | Sequence 12, Appl  |
| 31 | 47 | 100.0 | 609 | 13 | US-10-153-064-7   | Sequence 7, Appl   |
| 32 | 47 | 100.0 | 609 | 14 | US-10-153-604B-7  | Sequence 7, Appl   |
| 33 | 47 | 100.0 | 609 | 14 | US-10-365-623-23  | Sequence 23, Appl  |
| 34 | 47 | 100.0 | 610 | 9  | US-09-984-186-2   | Sequence 2, Appl   |
| 35 | 47 | 100.0 | 610 | 14 | US-10-237-667-2   | Sequence 2, Appl   |
| 36 | 47 | 100.0 | 610 | 14 | US-10-237-708-2   | Sequence 2, Appl   |
| 37 | 47 | 100.0 | 610 | 14 | US-10-237-866-2   | Sequence 2, Appl   |
| 38 | 47 | 100.0 | 610 | 14 | US-10-237-871-2   | Sequence 2, Appl   |
| 39 | 47 | 100.0 | 610 | 14 | US-10-237-624-2   | Sequence 2, Appl   |
| 40 | 47 | 100.0 | 616 | 12 | US-10-433-108-13  | Sequence 13, Appl  |
| 41 | 47 | 100.0 | 624 | 12 | US-10-433-108-16  | Sequence 16, Appl  |
| 42 | 47 | 100.0 | 631 | 12 | US-10-433-108-14  | Sequence 14, Appl  |
| 43 | 47 | 100.0 | 640 | 12 | US-10-433-108-15  | Sequence 15, Appl  |
| 44 | 47 | 100.0 | 640 | 12 | US-10-433-108-17  | Sequence 17, Appl  |
| 45 | 47 | 100.0 | 651 | 13 | US-10-153-064-133 | Sequence 133, Appl |

#### ALIGNMENTS

#### RESULT 1

US-10-296-115-799  
; Sequence 799, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hysq Inc  
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478  
; SEQ ID NO 799  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(114)  
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3  
US-10-296-115-799

Query Match 100.0%; Score 47; DB 12; Length 114;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9  
DB 79 EKPILLEKSH 87

#### RESULT 2

US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1

Query Match 100.0%; Score 47; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288

RESULT 10

US-09-833-245-18

; Sequence 18, Application US/09833245

; Publication No. US20040010134A1

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Albumin Fusion Proteins

; FILE REFERENCE: PF546PCT

; CURRENT APPLICATION NUMBER: US/09/833,245

; PRIOR FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229,358

; PRIOR FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: 60/256,931

; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199,384

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 2267

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 18

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-833-245-18

Query Match 100.0%; Score 47; DB 11; Length 585;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288

RESULT 11

US-10-424-999-11

; Sequence 11, Application US/10424999

; Publication No. US20040052810A1

; GENERAL INFORMATION:

; APPLICANT: Nesbitt, Mark

; APPLICANT: Cameron, Beatrice

; APPLICANT: Blanche, Francis

; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for

; FILE REFERENCE: ST01027-A

; CURRENT APPLICATION NUMBER: US/10/424,999

; PRIOR FILING DATE: 2003-04-29

; PRIOR APPLICATION NUMBER: 10/233,675

; PRIOR FILING DATE: 2002-09-04

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 11

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Fusion protein human abrogen

US-10-424-999-11

Query Match 100.0%; Score 47; DB 12; Length 585;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288

RESULT 12

US-10-425-000-31

; Sequence 31, Application US/10425000

; Publication No. US20040052777A1

; GENERAL INFORMATION:

; APPLICANT: Nesbitt, Mark

; APPLICANT: Cameron, Beatrice

; APPLICANT: Blanche, Francis

; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit

; FILE REFERENCE: ST01027-B

; CURRENT APPLICATION NUMBER: US/10/425,000

; PRIOR FILING DATE: 2003-04-29

; PRIOR APPLICATION NUMBER: 10/233,675

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 31

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Human derived fusion protein

US-10-425-000-31

Query Match

Best Local Similarity

Matches

9; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288

RESULT 13

US-10-433-108-34

; Sequence 34, Application US/10433108

; Publication No. US20040053370A1

; GENERAL INFORMATION:

; APPLICANT: Eli Lilly and Company

; TITLE OF INVENTION: GIP-1 FUSION PROTEINS

; FILE REFERENCE: X-13991

; CURRENT APPLICATION NUMBER: US/10/433,108

; PRIOR FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: US 60/251,954

; PRIOR FILING DATE: 2000-06-12

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 34

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-433-108-34

Query Match

Best Local Similarity

Matches

9; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288

RESULT 14

US-10-153-064-5

; Sequence 5, Application US/10153064

; Publication No. US20020142814A1

; GENERAL INFORMATION:

; APPLICANT: Bell et al.

```
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
```

```
Query Match      100.0%; Score 47; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EKPILLEKSH 9
Db 280 EKPILLEKSH 288
```

# RESULT 15

```
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5
```

```
Query Match      100.0%; Score 47; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EKPILLEKSH 9
Db 280 EKPILLEKSH 288
```

Search completed: April 19, 2004, 12:54:59  
Job time : 3.78947 secs



OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_439\_447

Perfect score: 51

Sequence: 1 KHPEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriaph.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description          |
|------------|-------|-------------|--------|-----------|----------------------|
| 1          | 51    | 100.0       | 396    | 4 Q81UK7  | Q81UK7 homo sapien   |
| 2          | 51    | 100.0       | 417    | 4 Q86YG0  | Q86YG0 homo sapien   |
| 3          | 39    | 76.5        | 125    | 16 Q8FDR1 | Q8FDR1 escherichia   |
| 4          | 39    | 76.5        | 125    | 16 Q8FDR1 | Q8FDR1 escherichia   |
| 5          | 39    | 76.5        | 254    | 2 Q9RLC5  | Q9RLC5 pseudomonas   |
| 6          | 39    | 76.5        | 352    | 16 Q91618 | Q91618 pseudomonas   |
| 7          | 39    | 76.5        | 352    | 16 Q8QX2  | Q8QX2 pseudomonas    |
| 8          | 39    | 76.5        | 352    | 16 Q8A98  | Q8A98 pseudomonas    |
| 9          | 38    | 74.5        | 55     | 16 Q92S6  | Q92S6 streptomyces   |
| 10         | 38    | 74.5        | 98     | 5 Q9YX6   | Q9YX6 caenorhabditis |
| 11         | 37    | 72.5        | 415    | 10 Q84VX7 | Q84VX7 arabidopsis   |
| 12         | 37    | 72.5        | 765    | 16 Q89KX8 | Q89KX8 bradyrhizob   |
| 13         | 37    | 72.5        | 1036   | 10 Q9C8K6 | Q9C8K6 arabidopsis   |
| 14         | 36    | 70.6        | 122    | 9 Q84MS7  | Q84MS7 bacteriophage |
| 15         | 36    | 70.6        | 224    | 16 Q8A5V9 | Q8A5V9 bacterioides  |
| 16         | 36    | 70.6        | 267    | 16 Q67673 | Q67673 aquifex aeo   |

|    |    |      |      |           |                     |
|----|----|------|------|-----------|---------------------|
| 17 | 36 | 70.6 | 321  | 11 Q8C846 | Q8C846 mus musculus |
| 18 | 36 | 70.6 | 375  | 4 Q96SY8  | Q96SY8 homo sapien  |
| 19 | 36 | 70.6 | 412  | 4 Q9NWF5  | Q9NWF5 homo sapien  |
| 20 | 36 | 70.6 | 527  | 4 Q96T01  | Q96T01 homo sapien  |
| 21 | 36 | 70.6 | 613  | 4 Q96SM7  | Q96SM7 homo sapien  |
| 22 | 36 | 70.6 | 754  | 4 Q9NC65  | Q9NC65 homo sapien  |
| 23 | 36 | 70.6 | 1485 | 4 Q9UL15  | Q9UL15 homo sapien  |
| 24 | 36 | 70.6 | 1486 | 4 Q8TF02  | Q8TF02 homo sapien  |
| 25 | 36 | 70.6 | 1644 | 4 Q9NUD8  | Q9NUD8 homo sapien  |
| 26 | 36 | 70.6 | 1685 | 4 Q96T89  | Q96T89 homo sapien  |
| 27 | 36 | 70.6 | 1740 | 4 Q8TEVO  | Q8TEVO homo sapien  |
| 28 | 36 | 70.6 | 1758 | 11 Q8CHN9 | Q8CHN9 rattus norv  |
| 29 | 36 | 70.6 | 1902 | 11 Q925Q1 | Q925Q1 mus musculus |
| 30 | 36 | 70.6 | 1939 | 4 Q91Y33  | Q91Y33 homo sapien  |
| 31 | 36 | 70.6 | 1957 | 4 Q81ZY8  | Q81ZY8 homo sapien  |
| 32 | 36 | 70.6 | 1999 | 4 Q8NFD6  | Q8NFD6 homo sapien  |
| 33 | 36 | 70.6 | 2165 | 4 Q8NFD5  | Q8NFD5 homo sapien  |
| 34 | 36 | 70.6 | 2285 | 4 Q9HBJ5  | Q9HBJ5 homo sapien  |
| 35 | 35 | 68.6 | 68   | 16 Q8A7F9 | Q8A7F9 bacterioides |
| 36 | 35 | 68.6 | 84   | 16 P96913 | P96913 mycobacteri  |
| 37 | 35 | 68.6 | 84   | 16 Q7UL14 | Q7UL14 mycobacteri  |
| 38 | 35 | 68.6 | 125  | 2 Q8VR61  | Q8VR61 escherichia  |
| 39 | 35 | 68.6 | 125  | 2 Q83W73  | Q83W73 escherichia  |
| 40 | 35 | 68.6 | 125  | 2 Q9VQ03  | Q9VQ03 escherichia  |
| 41 | 35 | 68.6 | 125  | 2 Q9AL42  | Q9AL42 shigella fl  |
| 42 | 35 | 68.6 | 125  | 16 Q8FAS7 | Q8FAS7 escherichia  |
| 43 | 35 | 68.6 | 125  | 15 Q83JN8 | Q83JN8 shigella fl  |
| 44 | 35 | 68.6 | 153  | 13 Q8QK3  | Q8QK3 fugu rubrip   |
| 45 | 35 | 68.6 | 182  | 5 Q8MZY6  | Q8MZY6 mamestra br  |

#### ALIGNMENTS

##### RESULT 1

Q81UK7 PRELIMINARY, PRT; 396 AA.

AC Q81UK7; TREMBLrel. 23, Created

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TremBLrel. 25, Last annotation update)

DE Similar to serum albumin precursor.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1] SEQUENCE FROM N.A.

RP TISSUE=Liver;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC035969; AAH35969.1; -

DR GO; GO:0005615; C:extracellular space; IEA.

DR GO; GO:0005386; F:carrier activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000264; Serum\_albumin.

DR Pfam; PF00273; transport\_prot; 2.

DR PRINTS; PR00802; SERUMALBUMIN.

DR SMART; SM00103; ALBUMIN; 2.

DR PROSITE; PS00212; ALBUMIN; 2.

SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match Best Local Similarity 100.0%; Score 51; DB 4; Length 396;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9

Db 250 KHPEAKRMP 258

##### RESULT 2

Q86YG0

Q86YGO PRELIMINARY; PRT; 417 AA.  
AC Q86YGO;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to alpha-fetoprotein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041789; AA41789.1;  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot.; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR PRODOM; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 417 AA; 47360 MW; 16E764833BEF4E8D CRC64;

Query Match 100.0%; Score 51; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KHPEAKR 9  
Db 271 KHPEAKRMP 279

RESULT 3  
Q8FDR1 PRELIMINARY; PRT; 125 AA.  
AC Q8FDR1  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein yeev.  
GN C3677.  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
DR EMBL; AE016766; AAN82125.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 125 AA; 14014 MW; 26AB67E398121996 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 125;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KHPEAKR 7  
Db 119 KHPEAKR 125

RESULT 4  
Q8FYS8 PRELIMINARY; PRT; 125 AA.  
AC Q8FYS8  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein yeev.  
GN C4576.  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
DR EMBL; AE016769; AAN83010.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 125 AA; 13964 MW; 473BED42C8398644 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 125;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KHPEAKR 7  
Db 119 KHPEAKR 125

RESULT 5  
Q9RLC5 PRELIMINARY; PRT; 254 AA.  
AC Q9RLC5  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Biotin synthetase (EC 2.8.1.6) (Fragment).  
GN B10B.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JM300;  
RA Graupner S., Wackernagel W.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ249742; CAB56476.1;  
DR GO; GO:0004076; F:biotin synthase activity; IEA.  
DR GO; GO:0016740; F:transferrase activity; IEA.  
DR GO; GO:0009102; P:biotin biosynthesis; IEA.  
DR InterPro; IPR002684; Biotin\_synth.  
DR InterPro; IPR006638; Elp3.  
DR InterPro; IPR007197; Radical SAM.  
DR Pfam; PF04055; Radical SAM; 1.  
DR SMART; SM00729; Elp3; 1.  
DR TIGRFAMs; TIGR00433; bioB; 1.  
KW Transferrase.  
SQ SEQUENCE 254 AA; 27771 MW; F512DF0FEFC12F67 CRC64;

Query Match 76.5%; Score 39; DB 2; Length 254;  
Best Local Similarity 77.8%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY      1 KHPEAKRMP 9
Db      109 KHPSAKDMP 117

RESULT 6
Q91618 ID Q91618 PRELIMINARY; PRT; 352 AA.
AC Q91618;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Biotin synthase.
GN BIOC OR PA0500.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Raizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL; AE004487; AAG03889.1; -.
DR PIR; F83582; F83582.
DR GO; GO:0004076; F:biotin synthase activity; IEA.
DR GO; GO:0009102; F:biotin biosynthesis; IEA.
DR InterPro; IPR002684; Biotin_synth.
DR InterPro; IPR006638; B1p3.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00433; bioB; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 39113 MW; E46491B67DDF369A CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      109 KHPSAKDMP 117

RESULT 7
Q98QX2 ID Q98QX2 PRELIMINARY; PRT; 352 AA.
AC Q98QX2;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Biotin synthetase.
GN BIOC OR PF0362.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,

```

```

Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016775; AAN65993.1; -.
DR TIGR; PP0362; -.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 38950 MW; 7EAC944861D3553F CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      109 KHPSAKDMP 117

RESULT 8
Q88A98 ID Q88A98 PRELIMINARY; PRT; 352 AA.
AC Q88A98;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Biotin synthetase.
GN BIOC OR PSPT00494.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidse T.,
RA White O., Fraser C., Collmer A.;
RT "Complete sequence of Pseudomonas syringae."
RT Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AE016857; AAC54038.1; -.
DR TIGR; PSPT00494; -.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 38610 MW; B2AB050D00F03F50 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      109 KHPSAKDMP 117

RESULT 9
Q9S2S6 ID Q9S2S6 PRELIMINARY; PRT; 55 AA.
AC Q9S2S6;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Putative small hydrophilic protein.
GN SCO2063 OR SC4G6.32.
OS Streptomyces coelicolor.

```

```

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2) / X145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL039111; CAB51455.1; -
DR FIR; T35092; T35092.
KW Complete proteome.
SQ SEQUENCE 55 AA; 6247 MW; 2F34EDC4E906354F CRC64;

Query Match 74.5%; Score 38; DB 16; Length 55;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
DB 11 KHPQAKRSP 19
|||||
|:|:|:|

RESULT 10
Q9GYG6 PRELIMINARY; PRT; 98 AA.
AC Q9GYG6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN W01C8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan M.;
RT "The sequence of C. elegans cosmid W01C8.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U41508; AAG00023.1; -
DR FIR; T26046; T26046.
DR WormPep; W01C8.1; CE05029.
KW Hypothetical protein.
SQ SEQUENCE 98 AA; 11596 MW; D65B2FFBF9FB859 CRC64;

Query Match 74.5%; Score 38; DB 5; Length 98;

```

```

Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HPEAKRMP 9
DB 13 HPEEKRLP 20
|||||
|:|:|:|

RESULT 11
Q84VX7 PRELIMINARY; PRT; 415 AA.
AC Q84VX7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Atg51540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RX Chan R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kaniya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BT004632; AAO42878.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR006652; Kelch rep.
DR Pfam; PF01344; Kelch; 4.
DR PROSITE; PS00344; MYB_2; 1.
SQ SEQUENCE 415 AA; 46747 MW; 2024DEEC55F82E08 CRC64;

Query Match 72.5%; Score 37; DB 10; Length 415;
Best Local Similarity 75.0%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 8
DB 338 KHPETKKM 345
|||||
|:|:|:|

RESULT 12
Q89KX8 PRELIMINARY; PRT; 765 AA.
AC Q89KX8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BIR4770 protein.
GN BIR4770.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RX MEDLINE=22484999; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Teurloka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";

```

Search completed: April 19, 2004, 12:00:13

Job time : 5.47368 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 5.29363 Seconds

(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_439\_447

Perfect score: 51

Sequence: 1 KXPEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 51    | 100.0       | 106    | 4 AAO04431 | Aao04431 Human pol |
| 2          | 51    | 100.0       | 119    | 4 AAO02645 | Aao02645 Human pol |
| 3          | 51    | 100.0       | 133    | 4 AAO03606 | Aao03606 Human pol |
| 4          | 51    | 100.0       | 151    | 4 AAU33085 | Aau33085 Novel hum |
| 5          | 51    | 100.0       | 187    | 4 AAU27873 | Aau27873 Human con |
| 6          | 51    | 100.0       | 228    | 3 AAY83949 | Aay83949 Yeast cod |
| 7          | 51    | 100.0       | 245    | 4 AAU33081 | Aau33081 Novel hum |
| 8          | 51    | 100.0       | 386    | 4 AAU33074 | Aau33074 Novel hum |
| 9          | 51    | 100.0       | 401    | 4 AAU23876 | Aau23876 Novel hum |
| 10         | 51    | 100.0       | 507    | 4 AAU33286 | Aau33286 Novel hum |
| 11         | 51    | 100.0       | 507    | 4 AAU32994 | Aau32994 Novel hum |
| 12         | 51    | 100.0       | 550    | 4 AAU29877 | Aau29877 Novel hum |
| 13         | 51    | 100.0       | 584    | 6 ABG72381 | Abg72381 Mature hu |
| 14         | 51    | 100.0       | 585    | 1 AAP93344 | Aap93344 Sequence  |
| 15         | 51    | 100.0       | 585    | 1 AAP90388 | Aap90388 Mature hu |
| 16         | 51    | 100.0       | 585    | 2 AAR03318 | Aar03318 Human ser |
| 17         | 51    | 100.0       | 585    | 2 AAR08457 | Aar08457 Human ser |
| 18         | 51    | 100.0       | 585    | 2 AAR26207 | Aar26207 Human ser |
| 19         | 51    | 100.0       | 585    | 2 AAR26362 | Aar26362 Synthetic |
| 20         | 51    | 100.0       | 585    | 2 AAR20029 | Aar20029 Human ser |
| 21         | 51    | 100.0       | 585    | 2 AAR80301 | Aar80301 Human ser |
| 22         | 51    | 100.0       | 585    | 2 AAO20111 | Aao20111 HSA prote |
| 23         | 51    | 100.0       | 585    | 2 AAWS9841 | Aaw59841 Mature pr |
| 24         | 51    | 100.0       | 585    | 3 AAY84873 | Aay84873 Amino aci |
| 25         | 51    | 100.0       | 585    | 3 AAY83946 | Aay83946 Yeast cod |

#### ALIGNMENTS

RESULT 1

AAO04431

ID AAO04431 standard; protein; 106 AA.

XX AC AAO04431;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 18323.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX PS WPI; 2001-514838/56.

XX DR N-PSDB; AAI84362.

XX XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 18323; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulatory activity, tissue growth factor activity, immunomodulatory activity and activity/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and

Aam52567 Mature hu  
Aae12417 Human alb  
Aae12403 Human alb  
Aae13129 Human alb  
Aae13135 Human alb  
Aae13111 Human alb  
Aae13399 Human alb  
Aab79006 Human mat  
Aae05878 Human ser  
Aau75220 Mature fo  
Abj00986 B lymphoc  
Abg63321 Human ser  
Abg33847 Human B L  
Abg71291 Glycosyla  
Abg72380 Mature hu  
Abr55695 Human alb  
Aae30936 Human alb  
Aap98782 Mature hu  
Abr42606 Human ser  
Adc16767 Human ser

26 51 100.0 585 4 AAM52567  
27 51 100.0 585 4 AAE12417  
28 51 100.0 585 4 AAE12403  
29 51 100.0 585 4 AAE13129  
30 51 100.0 585 4 AAE13135  
31 51 100.0 585 4 AAE13111  
32 51 100.0 585 4 AAE13399  
33 51 100.0 585 4 AAB79006  
34 51 100.0 585 4 AAE05878  
35 51 100.0 585 5 AAU75220  
36 51 100.0 585 5 ABJ00986  
37 51 100.0 585 5 ABG63321  
38 51 100.0 585 5 ABG33847  
39 51 100.0 585 5 ABG71291  
40 51 100.0 585 6 ABG72380  
41 51 100.0 585 6 ABR55695  
42 51 100.0 585 6 AAE30936  
43 51 100.0 585 6 AAP98782  
44 51 100.0 585 7 ABR42606  
45 51 100.0 585 7 ADC16767

CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Gaps 0;  
Matches 9; Conservative 0; Indels 0;

QY 1 KHPEAKRMP 9  
DB 65 KHPEAKRMP 73

## RESULT 2

AAO02645  
ID AAO02645 standard; protein; 119 AA.

XX AC AAO02645;

XX XX 06-NOV-2001 (first entry)

DE DE Human polypeptide SEQ ID NO 16537.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX XX 07-SEP-2001.

XX XX 26-FEB-2001; 2001WO-US004927.

XX XX 28-FEB-2000; 2000US-00515126.

XX XX 18-MAY-2000; 2000US-00577409.

XX XX (HYSE-) HYSEQ INC.

XX XX Tang YT, Liu C, Drmanac RT;

XX XX WPI; 2001-514838/56.

XX XX N-PSDB; AA182576.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 16537; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 119 AA;

Query Match 100.0%; Score 51; DB 4; Length 119;  
Best Local Similarity 100.0%; Pred. No. 0.14; Mismatches 0; Gaps 0;  
Matches 9; Conservative 0; Indels 0;

QY 1 KHPEAKRMP 9  
DB 46 KHPEAKRMP 54

## RESULT 3

AAO03606  
ID AAO03606 standard; protein; 133 AA.

XX AC AAO03606;

XX XX 06-NOV-2001 (first entry)

DE DE Human polypeptide SEQ ID NO 17498.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX XX 07-SEP-2001.

XX XX 26-FEB-2001; 2001WO-US004927.

XX XX 28-FEB-2000; 2000US-00515126.

XX XX 18-MAY-2000; 2000US-00577409;

XX XX (HYSE-) HYSEQ INC.

XX XX Tang YT, Liu C, Drmanac RT;

XX XX WPI; 2001-514838/56.

XX XX N-PSDB; AA183537.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 17498; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 133 AA;

Query Match 100.0%; Score 51; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Gaps 0;  
Matches 9; Conservative 0; Indels 0;

QY 1 KHPEAKRMP 9  
DB 78 KHPEAKRMP 86

## RESULT 4

AAU33085  
ID AAU33085 standard; protein; 151 AA.

XX AC AAU33085;



XX 18-DEC-2001 (first entry)  
 XX Novel human secreted protein #3576.  
 XX Human; vaccination; Gene therapy; nutritional supplement;  
 XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 OS Homo sapiens.  
 XX WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008656.  
 XX 18-APR-2000; 2000US-00552929.  
 XX 26-JAN-2001; 2001US-00770160.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 XX vaccination, testing and therapy.  
 XX Claim 20; Page 705; 765pp; English.  
 XX The invention relates to novel human secreted polypeptides. The  
 XX polypeptides and antibodies to the polypeptides are useful for  
 XX determining the presence of or predisposition to a disease associated  
 XX with altered levels of polypeptide. The polypeptides are also useful for  
 XX identifying agents (agonists and antagonists) that bind to them. Cells  
 XX expressing the proteins are useful for identifying a therapeutic agent  
 XX for use in treatment of a pathology related to aberrant expression or  
 XX physiological interactions of the polypeptide. Vectors comprising the  
 XX nucleic acids encoding the polypeptides and cells genetically engineered  
 XX to express them are also useful for producing the proteins. The proteins  
 XX are useful in genetic vaccination, testing and therapy, and can be used  
 XX as nutritional supplements. They may be used to increase stem cell  
 XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 XX and/or nerve tissue growth or regeneration; immune suppression and/or  
 XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 XX AAU29510-AAU3304 represent the amino acid sequences of novel human  
 XX secreted proteins of the invention  
 XX Sequence 151 AA;  
 Query Match 100.0%; Score 51; DB 4; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KHPKAKRMP 9  
 Db 12 KHPKAKRMP 20  
 RESULT 5  
 AAU27873  
 ID AAU27873 standard; protein; 187 AA.  
 AC AAU27873;  
 XX 18-DEC-2001 (first entry)  
 XX Human contig polypeptide sequence #26.  
 XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
 XX mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
 XX cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;

KW nervous system disorder; inflammatory disorder; cell differentiation;  
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
 KW cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;  
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;  
 KW immunostimulant; analgesic; gene therapy.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX WO200164834-A2.  
 XX 07-SEP-2001.  
 XX 26-FEB-2001; 2001WO-US004926.  
 XX 28-FEB-2000; 2000US-00515126.  
 XX 18-MAY-2000; 2000US-00577409.  
 XX 17-JUN-2000; 2000US-00597707.  
 XX 14-JUL-2000; 2000US-00616807.  
 XX 19-SEP-2000; 2000US-00664641.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 XX Drmanac R;  
 XX WPI; 2001-589862/66.  
 XX N-PSDB; AAS44773.  
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
 XX prepared from various human tissues, for diagnosis, treatment of cancer,  
 XX neurological, inflammatory disorders and for use in arrays for detection.  
 XX Claim 10; Page 127; 153pp; English.  
 XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig  
 XX polypeptides of the invention. The proteins and their associated DNA  
 XX sequences are useful for the treatment, diagnosis and prevention of  
 XX various types of disorder in a mammalian subject such as a human, dog,  
 XX monkey, mouse, hamster or rat. The disorders include cancers such as  
 XX leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 XX multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 XX disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 XX chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 XX Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 XX bowel disease. The sequences exhibit activity relating to angiogenesis,  
 XX cell proliferation, cell differentiation, stem cell growth factor,  
 XX activin or inhibin. Therefore, they can be used to manipulate stem cells  
 XX in culture to give rise to neuroepithelial cells that can be used to  
 XX augment or replace cells damaged by illness, accidental damage or genetic  
 XX disorders. The sequences may also be used for regeneration of bone,  
 XX cartilage, tendons and ligaments and in tissue repair and burn healing.  
 XX Note: Some sequences for this patent did not form part of the printed  
 XX specification, but were obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 187 AA;  
 Query Match 100.0%; Score 51; DB 4; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KHPKAKRMP 9  
 Db 75 KHPKAKRMP 84  
 RESULT 6

AAV83949  
ID AAY83949 standard; protein; 228 AA.  
XX  
AC AAY83949;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Yeast codon-biased recombinant HSA protein fragment HSA-III.  
XX  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN CN1239103-A.  
XX  
PD 22-DEC-1999.  
XX  
PF 17-JUN-1998; 98CN-00102506.  
XX  
PR 17-JUN-1998; 98CN-00102506.  
XX  
PA (HAJJI) HAJJI BIOENGINEERING CO LTD.  
XX  
PI Li S, Lu D;  
XX  
DR WPI; 2000-351198/31.  
DR N-PSDB; AAA10094.  
XX  
PT Process for preparing recombinant human serum albumin comprising yeast  
PT biased sex codons - uses a recombinant DNA technique.  
XX  
PS Example 1; Fig 7; 44pp; Chinese.  
XX  
CC The method relates to a method of recombinantly producing human serum  
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
CC a yeast codon bias. The complete HSA gene (AA10091) was generated as  
CC three synthetic fragments (AA10092-AA10094) joined by recombinant DNA  
CC technology. Each HSA fragment was synthesised from overlapping  
CC oligonucleotide fragments that were extended. This sequence represents  
CC the sequence of the HSA fragment HSA-III encoded by the human gene with a  
CC yeast codon bias. The invention also covers a recombinant expression  
CC vector, yeast host cells carrying the recombinant expression vector and  
CC the process for producing human serum albumin in the yeast host cell,  
CC especially in secretory mode  
XX  
SQ Sequence 228 AA;  
Query Match 100.0%; Score 51; DB 3; Length 228;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHEPAKMP 9  
DB 82 KHEPAKMP 90  
RESULT 7  
AAU33081  
ID AAU33081 standard; protein; 245 AA.  
XX  
AC AAU33081;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3572.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.

XX WO200179449-A2.  
PN  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR 26-JAN-2001; 2001US-00770160.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Drmanac RT;  
PI  
XX  
XX WPI; 2001-611725/70.  
DR  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
PT  
XX  
XX Claim 20; Page 705; 765pp; English.  
PS  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 245 AA;  
Query Match 100.0%; Score 51; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHEPAKMP 9  
DB 167 KHEPAKMP 175  
RESULT 8  
AAU33074  
ID AAU33074 standard; protein; 386 AA.  
XX  
AC AAU33074;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3565.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR

```
PR 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 704; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 386 AA;
Query Match 100.0%; Score 51; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHEPAKRM 9
DB 159 KHEPAKRM 167
|||||
159 KHEPAKRM 167

RESULT 9
AAU29876
ID AAU29876 standard; protein; 401 AA.
AC AAU29876;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #367.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 386 AA;
Query Match 100.0%; Score 51; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHEPAKRM 9
DB 159 KHEPAKRM 167
|||||
159 KHEPAKRM 167

RESULT 9
AAU29876
ID AAU29876 standard; protein; 401 AA.
AC AAU29876;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3777.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 401 AA;
Query Match 100.0%; Score 51; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHEPAKRM 9
DB 255 KHEPAKRM 263
|||||
255 KHEPAKRM 263

RESULT 10
AAU33286
ID AAU33286 standard; protein; 507 AA.
AC AAU33286;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3777.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 401 AA;
Query Match 100.0%; Score 51; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHEPAKRM 9
DB 255 KHEPAKRM 263
|||||
255 KHEPAKRM 263
```

CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 507 AA;

Query Match 100.0%; Score 51; DB 4; Length 507;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
|||||  
Db 215 KHPEAKRMP 223

RESULT 11  
AAU32994  
ID AAU32994 standard; protein; 507 AA.  
XX  
AC AAU32994;  
XX

DT 18-DEC-2001 (first entry)  
XX

DE Novel human secreted protein #3485.  
XX

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX

OS Homo sapiens.  
XX

PN WO200179449-A2.  
XX

PD 25-OCT-2001.  
XX

PF 16-APR-2001; 2001WO-US008656.  
XX

PR 18-APR-2000; 2000US-00552929.  
XX

PR 26-JAN-2001; 2001US-00770160.  
XX

XX (HYSE-) HYSEQ INC.  
PA

XX Tang YT, Liu C, Drmanac RT;  
PI

XX WPI; 2001-611725/70.  
DR

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX

PS Claim 20; Page 696; 765pp; English.  
XX

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 507 AA;

Query Match 100.0%; Score 51; DB 4; Length 507;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
|||||  
Db 215 KHPEAKRMP 223

RESULT 12  
AAU29877

ID AAU29877 standard; protein; 550 AA.  
XX

AC AAU29877;  
XX

DT 18-DEC-2001 (first entry)  
XX

DE Novel human secreted protein #368.  
XX

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX

OS Homo sapiens.  
XX

PN WO200179449-A2.  
XX

PD 25-OCT-2001.  
XX

PF 16-APR-2001; 2001WO-US008656.  
XX

PR 18-APR-2000; 2000US-00552929.  
XX

PR 26-JAN-2001; 2001US-00770160.  
XX

XX (HYSE-) HYSEQ INC.  
PA

XX Tang YT, Liu C, Drmanac RT;  
PI

XX WPI; 2001-611725/70.  
DR

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX

PS Claim 20; Page 206; 765pp; English.  
XX

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX

SQ Sequence 550 AA;

```

Query Match      100.0%; Score 51; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      404 KHPEAKRMP 412

RESULT 13
ABG72381
ID ABG72381 standard; protein; 584 AA.
XX
AC ABG72381;
XX
DT 10-FEB-2003 (first entry)
DE
DE Mature human serum albumin #2.
XX
KW Human; serum albumin; HSA; cancer; cytostatic; breast cancer;
KW prostate cancer; anti-proliferative.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 241..242
FT /note= "Encoded by GTCCACAGC"
XX
PN US2002123080-A1.
XX
PD 05-SEP-2002.
XX
PF 14-AUG-2001; 2001US-00929552.
XX
PR 19-DEC-1996; 96US-00769746.
XX
PA (TUFT ) UNIV TUFTS.
XX
PI Sonnenschein C, Soto AM;
XX
DR WPI; 2003-066789/06.
DR N-PSDB; ABX13582.
XX
PT Testing human cancer cells, especially breast and prostate cancer cells,
PT by contacting cancer cells obtained from biopsy of a patient ex vivo with
PT antibody specific to human albumin receptor.
XX
PS Disclosure; Fig 2; 17pp; English.
XX
CC The invention relates to testing human cancer cells, comprising obtaining
CC cancer cells from the patient and contacting the cell ex vivo with an
CC antibody to the receptor for human albumin. The method is useful for
CC testing human cancer cells in particular breast and prostate cancer
CC cells. The present sequence is mature human serum albumin, HSA. The anti-
CC proliferative effect of HSA was assayed in an experiment included in the
CC specification
XX
SQ Sequence 584 AA;

Query Match      100.0%; Score 51; DB 6; Length 584;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      438 KHPEAKRMP 446

RESULT 14
AAP93344
ID AAP93344 standard; protein; 585 AA.
XX
AC AAP93344;
XX
DT 25-MAR-2003 (revised)
DT 23-JUN-1990 (first entry)
XX
DE Sequence of mature human serum albumin (HSA) as encoded by artificial
DE gene.
XX
KW Mature human serum albumin; artificial gene; oligonucleotide block;
KW hypobolaemia; shock; hypoalbuminaemia.
XX
OS Homo sapiens.
XX
PN EP308381-A.
XX
PD 22-MAR-1989.
XX
PF 13-SEP-1988; 88EP-00850299.
XX
PR 14-SEP-1987; 87SE-00003539.
XX
PA (SKAN-) SKANDIGEN AB.
PA (MAGY ) MTA SZEGEDI BIOLOG KOEZFONTI.
PA (VEPE-) VEPEX CONTRACTOR LTD.
XX
PI Aberg B, Simoncsits A, Kalan M, Csperpan I, Bajszar G;
XX
DR WPI; 1989-087749/12.
DR N-PSDB; AAN90937.
XX
PT Artificial gene coding for authentic human serum albumin - constructed on
PT the basis of codons most frequently used by chosen non-human host.
XX
PS Disclosure; pp. 11-16; 121pp; English.
XX
CC The synthetic gene was constructed by designing a nucleotide sequence in
CC which the codons which are most frequently used by the chosen non-human
CC host were selected. In this case, it is yeast cells (LL20; Leu2-3; 112,
CC His 3-11, 15). The synthetic HSA gene was assembled from 24
CC oligonucleotide blocks. HSA is used in therapy for the treatment of
CC hypovolaemia, shock and hypoalbuminaemia. It is also used as an additive
CC in perfusion liq. for extracorporeal circulation and as an experimental
CC antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR
CC -2003 to correct PI field.)
XX
SQ Sequence 585 AA;

Query Match      100.0%; Score 51; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      439 KHPEAKRMP 447

RESULT 15
AAP90388
ID AAP90388 standard; protein; 585 AA.
XX
AC AAP90388;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Mature human serum albumin polypeptide.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
PN EP322094-A.
```

XX 28-JUN-1989.  
PD  
XX  
XX 25-OCT-1988; 88EP-00310000.  
PF  
XX  
XX 30-OCT-1987; 87GB-00025529.  
PR  
XX  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA  
XX  
XX Balance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
PI  
XX  
XX WPI; 1989-186464/26.  
DR  
XX N-PSDB; AAN90128.  
DR  
XX  
XX New N-terminal fragments of human serum albumin - esp. useful as blood  
PT plasma expanders.  
PI  
XX  
XX Disclosure; Fig 2; 20pp; English.  
PS  
XX  
XX Mature protein of human serum albumin (see corresp. AAN90128). Used to  
CC make new N-terminal fragments which are used as plasma expanders, or as  
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-  
CC 2003 to correct FA field.) (Updated on 24-OCT-2003 to standardise OS  
CC field)  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 51; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.77;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHPEAKRMP 9  
Db 439 KHPEAKRMP 447  
|||||

Search completed: April 19, 2004, 11:51:21  
Job time : 6.29363 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_439\_447

Perfect score: 51

Sequence: 1 KHEAKMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 51    | 100.0       | 585    | 9  | US-09-929-552-2   |
| 2          | 51    | 100.0       | 585    | 10 | US-09-932-613-445 |
| 3          | 51    | 100.0       | 585    | 10 | US-09-984-010-26  |
| 4          | 51    | 100.0       | 585    | 10 | US-09-833-041-18  |
| 5          | 51    | 100.0       | 585    | 10 | US-09-833-117-18  |
| 6          | 51    | 100.0       | 585    | 10 | US-09-932-322-445 |
| 7          | 51    | 100.0       | 585    | 10 | US-09-832-501-18  |
| 8          | 51    | 100.0       | 585    | 11 | US-09-833-118-18  |
| 9          | 51    | 100.0       | 585    | 11 | US-09-833-245-18  |
| 10         | 51    | 100.0       | 585    | 12 | US-10-424-999-11  |
| 11         | 51    | 100.0       | 585    | 12 | US-10-425-000-31  |
| 12         | 51    | 100.0       | 585    | 12 | US-10-433-108-34  |
| 13         | 51    | 100.0       | 585    | 13 | US-10-153-064-5   |
| 14         | 51    | 100.0       | 585    | 14 | US-10-153-604A-5  |
| 15         | 51    | 100.0       | 585    | 14 | US-10-319-263-1   |

|    |    |       |     |    |                    |                    |
|----|----|-------|-----|----|--------------------|--------------------|
| 16 | 51 | 100.0 | 585 | 14 | US-10-319-263-2    | Sequence 2, Appli  |
| 17 | 51 | 100.0 | 585 | 14 | US-10-414-469-1    | Sequence 1, Appli  |
| 18 | 51 | 100.0 | 585 | 14 | US-10-414-469-2    | Sequence 2, Appli  |
| 19 | 51 | 100.0 | 585 | 14 | US-10-413-831-1    | Sequence 1, Appli  |
| 20 | 51 | 100.0 | 585 | 14 | US-10-413-831-2    | Sequence 2, Appli  |
| 21 | 51 | 100.0 | 585 | 15 | US-10-413-832-1    | Sequence 1, Appli  |
| 22 | 51 | 100.0 | 585 | 15 | US-10-413-832-2    | Sequence 2, Appli  |
| 23 | 51 | 100.0 | 585 | 15 | US-10-414-386-1    | Sequence 1, Appli  |
| 24 | 51 | 100.0 | 585 | 15 | US-10-414-386-2    | Sequence 2, Appli  |
| 25 | 51 | 100.0 | 585 | 15 | US-10-233-675A-11  | Sequence 11, Appli |
| 26 | 51 | 100.0 | 585 | 15 | US-10-462-262-26   | Sequence 26, Appli |
| 27 | 51 | 100.0 | 604 | 10 | US-09-984-010-7    | Sequence 7, Appli  |
| 28 | 51 | 100.0 | 609 | 10 | US-09-919-039-370  | Sequence 370, App  |
| 29 | 51 | 100.0 | 609 | 12 | US-10-609-346-12   | Sequence 12, Appli |
| 30 | 51 | 100.0 | 609 | 13 | US-10-153-064-7    | Sequence 7, Appli  |
| 31 | 51 | 100.0 | 609 | 14 | US-10-153-604A-7   | Sequence 7, Appli  |
| 32 | 51 | 100.0 | 609 | 14 | US-10-365-623-23   | Sequence 23, Appli |
| 33 | 51 | 100.0 | 610 | 9  | US-09-984-186-2    | Sequence 2, Appli  |
| 34 | 51 | 100.0 | 610 | 14 | US-10-237-667-2    | Sequence 2, Appli  |
| 35 | 51 | 100.0 | 610 | 14 | US-10-237-708-2    | Sequence 2, Appli  |
| 36 | 51 | 100.0 | 610 | 14 | US-10-237-866-2    | Sequence 2, Appli  |
| 37 | 51 | 100.0 | 610 | 14 | US-10-237-871-2    | Sequence 2, Appli  |
| 38 | 51 | 100.0 | 610 | 14 | US-10-237-824-2    | Sequence 2, Appli  |
| 39 | 51 | 100.0 | 616 | 12 | US-10-433-108-13   | Sequence 13, Appli |
| 40 | 51 | 100.0 | 624 | 12 | US-10-433-108-16   | Sequence 16, Appli |
| 41 | 51 | 100.0 | 631 | 12 | US-10-433-108-14   | Sequence 14, Appli |
| 42 | 51 | 100.0 | 640 | 12 | US-10-433-108-15   | Sequence 15, Appli |
| 43 | 51 | 100.0 | 640 | 12 | US-10-433-108-17   | Sequence 17, Appli |
| 44 | 51 | 100.0 | 651 | 13 | US-10-153-064-133  | Sequence 133, App  |
| 45 | 51 | 100.0 | 651 | 14 | US-10-153-604A-133 | Sequence 133, App  |

ALIGNMENTS

RESULT 1  
US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids





```
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 6
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, M. Daniel
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BlyS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932.322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832.501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833.118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 51; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833.245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18
```

Query Match 100.0%; Score 51; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
Db 439 KHPEAKRMP 447

RESULT 10  
US-10-424-999-11  
; Sequence 11, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbitt, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Fusion protein human abrogen  
US-10-424-999-11

Query Match 100.0%; Score 51; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
Db 439 KHPEAKRMP 447

RESULT 11  
US-10-425-000-31  
; Sequence 31, Application US/10425000  
; Publication No. US20040052777A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbitt, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit  
; TITLE OF INVENTION: Angiogenesis  
; FILE REFERENCE: ST01027-B  
; CURRENT APPLICATION NUMBER: US/10/425,000  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 31  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human derived fusion protein  
US-10-425-000-31

Query Match 100.0%; Score 51; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9

Db 439 KHPEAKRMP 447

RESULT 12  
US-10-433-108-34  
; Sequence 34, Application US/10433108  
; Publication No. US20040053370A1  
; GENERAL INFORMATION:  
; APPLICANT: Eli Lilly and Company  
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS  
; FILE REFERENCE: X-13991  
; CURRENT APPLICATION NUMBER: US/10/433,108  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: US 60/251,954  
; PRIOR FILING DATE: 2000-06-12  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-433-108-34

Query Match 100.0%; Score 51; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
Db 439 KHPEAKRMP 447

RESULT 13  
US-10-153-064-5  
; Sequence 5, Application US/10153064  
; Publication No. US20020142814A1  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-5

Query Match 100.0%; Score 51; DB 13; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9  
Db 439 KHPEAKRMP 447

RESULT 14  
US-10-153-604A-5  
; Sequence 5, Application US/10153604A  
; Publication No. US20030143191A1  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,604A  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212

```
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version.3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match      100.0%; Score 51; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 15
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1

Query Match      100.0%; Score 51; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

Search completed: April 19, 2004, 12:54:59
Job time : 3.78947 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.3795 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_280\_288

Perfect score: 47  
Sequence: 1 EKPLLEKSH 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/aaa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/2/aaa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/2/aaa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/2/aaa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/2/aaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/2/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                  | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 47    | 100.0       | 585    | 1 US-08-153-799-14  | Sequence 14, Appl  |
| 2          | 47    | 100.0       | 585    | 1 US-08-448-196A-3  | Sequence 3, Appl   |
| 3          | 47    | 100.0       | 585    | 2 US-08-984-176-1   | Sequence 1, Appl   |
| 4          | 47    | 100.0       | 585    | 2 US-08-702-572-2   | Sequence 2, Appl   |
| 5          | 47    | 100.0       | 585    | 3 US-08-769-746-2   | Sequence 2, Appl   |
| 6          | 47    | 100.0       | 585    | 4 US-10-153-064-5   | Sequence 5, Appl   |
| 7          | 47    | 100.0       | 609    | 1 US-08-222-619-3   | Sequence 3, Appl   |
| 8          | 47    | 100.0       | 609    | 4 US-08-433-037-4   | Sequence 4, Appl   |
| 9          | 47    | 100.0       | 609    | 4 US-08-897-956A-2  | Sequence 2, Appl   |
| 10         | 47    | 100.0       | 609    | 4 US-10-153-064-7   | Sequence 7, Appl   |
| 11         | 47    | 100.0       | 609    | 4 US-09-976-594-977 | Sequence 977, Appl |
| 12         | 47    | 100.0       | 609    | 5 PCT-US95-04075-3  | Sequence 3, Appl   |
| 13         | 47    | 100.0       | 610    | 2 US-08-797-689-2   | Sequence 2, Appl   |
| 14         | 47    | 100.0       | 610    | 4 US-09-984-186-2   | Sequence 2, Appl   |
| 15         | 47    | 100.0       | 651    | 4 US-10-153-064-133 | Sequence 133, Appl |
| 16         | 47    | 100.0       | 652    | 4 US-10-153-064-96  | Sequence 96, Appl  |
| 17         | 47    | 100.0       | 652    | 4 US-10-153-064-99  | Sequence 99, Appl  |
| 18         | 47    | 100.0       | 652    | 4 US-10-153-064-105 | Sequence 105, Appl |
| 19         | 47    | 100.0       | 652    | 4 US-10-153-064-132 | Sequence 132, Appl |
| 20         | 47    | 100.0       | 653    | 4 US-10-153-064-131 | Sequence 131, Appl |
| 21         | 47    | 100.0       | 656    | 4 US-10-153-064-130 | Sequence 130, Appl |
| 22         | 47    | 100.0       | 660    | 4 US-10-153-064-90  | Sequence 90, Appl  |
| 23         | 47    | 100.0       | 660    | 4 US-10-153-064-93  | Sequence 93, Appl  |
| 24         | 47    | 100.0       | 668    | 4 US-10-153-064-102 | Sequence 102, Appl |
| 25         | 47    | 100.0       | 676    | 4 US-10-153-064-95  | Sequence 95, Appl  |
| 26         | 47    | 100.0       | 676    | 4 US-10-153-064-98  | Sequence 98, Appl  |
| 27         | 47    | 100.0       | 676    | 4 US-10-153-064-104 | Sequence 104, Appl |

Sequence 127, App  
Sequence 129, App  
Sequence 125, App  
Sequence 123, App  
Sequence 92, Appl  
Sequence 101, App  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 89, Appl  
Sequence 1, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 5132, Ap

ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swops, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-19)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"  
US-08-153-799-14

Query Match 100.0%; Score 47; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9  
| | | | | | | | | |  
DB 280 EKPLEKSH 288

## RESULT 2

US-08-448-196A-3  
Sequence 3, Application US/08448196A  
Patent No. 5780594

GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C.  
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
TITLE OF INVENTION: RELATED PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NASA  
STREET: MARSHALL SPACE FLIGHT CENTER  
CITY: HUNTSVILLE  
STATE: ALABAMA  
COUNTRY: USA  
ZIP: 35812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,196A  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: BROAD JR., ROBERT L.  
REGISTRATION NUMBER: 18,757  
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
TELEPHONE: 205-544-0021  
TELEFAX: 205-544-0258  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-448-196A-3

Query Match 100.0%; Score 47; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9  
| | | | | | | | | |  
DB 280 EKPLEKSH 288

## RESULT 3

US-08-984-176-1  
Sequence 1, Application US/08984176  
Patent No. 5948609

GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C  
APPLICANT: HO, JOSEPH X  
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER  
FILE REFERENCE: 08/984,176  
CURRENT APPLICATION NUMBER: US/08/984,176  
CURRENT FILING DATE: 1997-12-03  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 1

LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-984-176-1

Query Match 100.0%; Score 47; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9  
| | | | | | | | | |  
DB 280 EKPLEKSH 288

## RESULT 4

US-08-702-572-2  
Sequence 2, Application US/08702572  
Patent No. 5965386

GENERAL INFORMATION:  
APPLICANT: Kerry-Williams, Sean M  
APPLICANT: Gilbert, Sarah C  
TITLE OF INVENTION: Yeast Strains and Modified Albumins  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Centeon L.L.C.  
STREET: 1020 First Avenue  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406-1310

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,572  
FILING DATE: 11-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 95/23857  
FILING DATE: 1-MAR-1995  
APPLICATION NUMBER: GB 9404270.2  
FILING DATE: 5-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Naomi Biswas  
REGISTRATION NUMBER: 38,384  
REFERENCE/DOCKET NUMBER: CS0114 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610/878/4294  
TELEFAX: 610/878/4221  
INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 595 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-702-572-2

Query Match 100.0%; Score 47; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 5  
US-08-769-746-2  
; Sequence 2, Application US/08769746  
; Patent No. 6274305  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; APPLICANT: Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/769,746  
; FILING DATE: 19-DEC-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-769-746-2

Query Match 100.0%; Score 47; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 6  
US-10-153-064-5  
; Sequence 5, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: P2556

; CURRENT APPLICATION NUMBER: US/10/153.064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-5

Query Match 100.0%; Score 47; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 7  
US-08-222-619-3  
; Sequence 3, Application US/08222619  
; Patent No. 5652352  
; GENERAL INFORMATION:  
; APPLICANT: Lichenstein, Henri  
; APPLICANT: Lyons, David  
; APPLICANT: Wurfel, Mark  
; APPLICANT: Wright, Samuel  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,619  
; FILING DATE:  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 304 EKPLLEKSH 312

RESULT 8  
US-08-433-037-4  
; Sequence 4, Application US/08433037  
; Patent No. 5707828  
; GENERAL INFORMATION:  
; APPLICANT: Sreekrishna, Kotikanyadan

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 595 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-702-572-2

Query Match 100.0%; Score 47; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 5  
US-08-769-746-2  
; Sequence 2, Application US/08769746  
; Patent No. 6274305  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; APPLICANT: Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/769,746  
; FILING DATE: 19-DEC-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-769-746-2

Query Match 100.0%; Score 47; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 6  
US-10-153-064-5  
; Sequence 5, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: P2556

; CURRENT APPLICATION NUMBER: US/10/153.064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-5

Query Match 100.0%; Score 47; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 280 EKPLLEKSH 288

RESULT 7  
US-08-222-619-3  
; Sequence 3, Application US/08222619  
; Patent No. 5652352  
; GENERAL INFORMATION:  
; APPLICANT: Lichenstein, Henri  
; APPLICANT: Lyons, David  
; APPLICANT: Wurfel, Mark  
; APPLICANT: Wright, Samuel  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,619  
; FILING DATE:  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
Db 304 EKPLLEKSH 312

RESULT 8  
US-08-433-037-4  
; Sequence 4, Application US/08433037  
; Patent No. 5707828  
; GENERAL INFORMATION:  
; APPLICANT: Sreekrishna, Kotikanyadan

APPLICANT: Barr, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Techopp, Jurg F.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,345  
REFERENCE/DOCKET NUMBER: 91082  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9  
DB 304 EKPLLEKSH 312

RESULT 9  
US-08-897-956A-2  
Sequence 2, Application US/08897956A  
Patent No. 6423512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 47; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
DB 304 EKPLLEKSH 312  
RESULT 10  
US-10-153-064-7  
Sequence 7, Application US/10153064  
Patent No. 6663485  
GENERAL INFORMATION:  
APPLICANT: Bell et al.  
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
FILE REFERENCE: PFS56  
CURRENT APPLICATION NUMBER: US/10/153,064  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 60/293,212  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 7  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 47; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9  
DB 304 EKPLLEKSH 312

RESULT 11  
US-09-976-594-977  
Sequence 977, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 977  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 089957CD1  
US-09-976-594-977

Query Match 100.0%; Score 47; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9  
DB 304 EKPLLEKSH 312

RESULT 12  
PCT-US95-04075-3  
Sequence 3, Application PC/TUS9504075  
GENERAL INFORMATION:  
APPLICANT: AMGEN INC.

; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
 ; TITLE OF INVENTION: Protein  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Amgen Center, Patent Operations/RR  
 ; STREET: 1840 DeHavilland Drive  
 ; CITY: Thousand Oaks  
 ; STATE: California  
 ; COUNTRY: U.S.  
 ; ZIP: 91320-1789  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04075  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 609 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; PCT-US95-04075-3

Query Match 100.0%; Score 47; DB 5; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
 DB 304 EKPLLEKSH 312

RESULT 13  
 US-08-797-689-2  
 ; Sequence 2, Application US/08797689  
 ; Patent No. 5876969  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fleer, Reinhard  
 ; APPLICANT: Fournier, Alain  
 ; APPLICANT: Guittion, Jean-Dominique  
 ; APPLICANT: Jung, Gerard  
 ; APPLICANT: Yeh, Patrice  
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
 ; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 ; STREET: 500 Arcola Road, 3C43  
 ; CITY: Collegeville  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19426  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: Macintosh  
 ; OPERATING SYSTEM: System 7.1  
 ; SOFTWARE: Word 5.1 (PatentIn)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/797,689  
 ; FILING DATE: 31-JAN-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/256,927  
 ; FILING DATE: 28-JUL-1994  
 ; APPLICATION NUMBER: FR 92/01064  
 ; FILING DATE: 31-JAN-1992  
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FR93/00085  
 ; FILING DATE: 28-JAN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Smith Ph.D., Julie K.  
 ; REGISTRATION NUMBER: P-38,619  
 ; REFERENCE/DOCKET NUMBER: ST92006-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (610) 454-3839  
 ; TELEFAX: (610) 454-3808  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 610 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-797-689-2

Query Match 100.0%; Score 47; DB 2; Length 610;  
 Best Local Similarity 100.0%; Pred. No. 0.35;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9  
 DB 304 EKPLLEKSH 312

RESULT 14  
 US-09-984-186-2  
 ; Sequence 2, Application US/09984186  
 ; Patent No. 6686179  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fleer, Reinhard  
 ; APPLICANT: Fournier, Alain  
 ; APPLICANT: Guittion, Jean-Dominique  
 ; APPLICANT: Jung, Gerard  
 ; APPLICANT: Yeh, Patrice  
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
 ; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
 ; CONTAINING SAID POLYPEPTIDES  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 ; STREET: 500 Arcola Road, 3C43  
 ; CITY: Collegeville  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19426  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: Macintosh  
 ; OPERATING SYSTEM: System 7.1  
 ; SOFTWARE: Word 5.1 (PatentIn)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/984,186  
 ; FILING DATE: 29-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/797,689  
 ; FILING DATE: 31-JAN-1997  
 ; APPLICATION NUMBER: US 08/256,927  
 ; FILING DATE: 28-JUL-1994  
 ; APPLICATION NUMBER: FR 92/01064  
 ; FILING DATE: 31-JAN-1992  
 ; APPLICATION NUMBER: PCT/FR93/00085  
 ; FILING DATE: 28-JAN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Smith Ph.D., Julie K.  
 ; REGISTRATION NUMBER: P-38,619  
 ; REFERENCE/DOCKET NUMBER: ST92006-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (610) 454-3839  
 ; TELEFAX: (610) 454-3808  
 ; INFORMATION FOR SEQ ID NO: 2:



```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 47; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9
Db 304 EKPILLEKSH 312

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 47; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9
Db 346 EKPILLEKSH 354

```

Search completed: April 19, 2004, 12:05:20  
Job time : 1.3795 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.853186 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_362\_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 38    | 100.0       | 600    | 2 A47391 | serum albumin prec |
| 2          | 38    | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 3          | 34    | 89.5        | 177    | 2 D71311 | hypothetical prote |
| 4          | 34    | 89.5        | 575    | 2 C70781 | probable fprB prot |
| 5          | 34    | 89.5        | 740    | 2 E69420 | hydrogenase expres |
| 6          | 33    | 86.8        | 238    | 2 H75392 | hypothetical prote |
| 7          | 33    | 86.8        | 308    | 2 F86147 | hypothetical prote |
| 8          | 33    | 86.8        | 345    | 2 AH3101 | IS5 family transpo |
| 9          | 33    | 86.8        | 351    | 2 B98185 | hypothetical prote |
| 10         | 33    | 86.8        | 414    | 1 C70859 | probable hexosyltr |
| 11         | 33    | 86.8        | 438    | 2 E87123 | probable transfera |
| 12         | 32    | 84.2        | 363    | 2 A95504 | X-pro aminopeptida |
| 13         | 32    | 84.2        | 421    | 2 AH3200 | hydroxypyruvate re |
| 14         | 31    | 81.6        | 105    | 2 C72631 | hypothetical prote |
| 15         | 31    | 81.6        | 238    | 1 S72623 | ribulose-5-phospha |
| 16         | 31    | 81.6        | 262    | 2 D71281 | probable pyroline  |
| 17         | 31    | 81.6        | 272    | 2 T32959 | hypothetical prote |
| 18         | 31    | 81.6        | 300    | 2 T34213 | hypothetical prote |
| 19         | 31    | 81.6        | 345    | 2 T25833 | hypothetical prote |
| 20         | 31    | 81.6        | 418    | 2 A96601 | hypothetical prote |
| 21         | 31    | 81.6        | 634    | 2 T49415 | hypothetical prote |
| 22         | 31    | 81.6        | 713    | 2 A28706 | calpastatin, cardi |
| 23         | 31    | 81.6        | 883    | 2 A84210 | hypothetical prote |
| 24         | 31    | 81.6        | 1316   | 2 D87145 | [beta], subunit of |
| 25         | 31    | 81.6        | 1316   | 2 S31146 | DNA-directed RNA p |
| 26         | 31    | 81.6        | 1316   | 2 G70535 | probable rpoC prot |
| 27         | 31    | 81.6        | 1863   | 1 A58881 | breat/ovarian can  |
| 28         | 30    | 78.9        | 152    | 2 C75544 | hypothetical prote |
| 29         | 30    | 78.9        | 232    | 2 G70738 | probable PPE prote |

## ALIGNMENTS

### RESULT 1

A47391

serum albumin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999

C:Accession: A47391

R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwolet, J.; Putnam, F

Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993

A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir

A:Reference number: A47391; MUID:93211971; PMID:8460152

A:Contents: B/B homozygote

A:Accession: A47391

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-600 <WAT>

A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295

A:Experimental source: liver

A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281)

C:Superfamily: serum albumin; serum albumin repeat homology

F:21-194/Domain: serum albumin repeat homology <SA1>

F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 38; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7

DB 378 AAADPHE 384

### RESULT 2

ABHUS

serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

N:Contains: kinetensin

C:Species: Homo sapiens (man)

C>Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text\_change 17-Mar-2000

C:Accession: A93743; A93936; J39427; I59286; I59313; G01747; S55314; A91420; S06422; S31

R:Law, R.M.; Aaelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu:

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col:

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K' 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; GB:L00133; NID:G285591; PIDN:CAA2:

R:Dugaiczyk, A.; Law, S.W.; dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936  
A;Molecule type: mRNA  
A;Residues: 1-120, 'G', 122-608 <DUG>  
A;Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590  
U;Rano, Y.; Watanabe, K.; Sakai, M.; Tamaki, I.  
J. Biol. Chem. 261, 3244-3251, 1986  
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A;Reference number: 139427; MUID:86140099; PMID:2419329  
A;Accession: 139427  
A;Status: translation not shown  
A;Molecule type: protein  
A;Residues: 1-26 <URA>  
A;Cross-references: GB:M13075; NID:9178330; PIDN:AA51688.1; PID:9553173  
R;Watkins, S.; Madison, J.; Galliano, M.; Minichioti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A;Reference number: 159286; MUID:94181575; PMID:8134387  
A;Accession: 159286  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 282-290, 'KSRFDLQ' <WAT>  
A;Cross-references: GB:S69192; NID:9546032; PIDN:AA830282.1; PID:9546033  
A;Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia  
R;Madison, J.; Galliano, M.; Watkins, S.; Minichioti, L.; Porta, F.; Rossi, A.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A;Reference number: 159313; MUID:94294404; PMID:8022807  
A;Accession: 159313  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 589-590, 'ALPRVKNLLQVLP' <MAD>  
A;Cross-references: GB:S70799; NID:9547231; PIDN:AA831177.1; PID:9547232  
A;Note: This frame-shift variant is designated albumin Bazzano; four additional variants  
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.  
Biochem. J. 308, 321-325, 1995  
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
A;Reference number: 159314; MUID:95275251; PMID:7755581  
A;Accession: 159314  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-120, 'G', 122-455 <VEN>  
A;Cross-references: EMBL:U22361; NID:9763428; PIDN:AAA64902.1; PID:9763431  
R;Edgewood, E.C.; George, P.M.; Feach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A;Title: Complete amino acid sequence of human serum albumin.  
A;Reference number: A91420; MUID:76187907; PMID:1225573  
A;Accession: A91420  
A;Molecule type: protein  
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
R;Roehr, U.; Spitteller, G.; Triptier, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A;Reference number: S06422  
A;Note: this paper is in German, with an English abstract  
A;Accession: S06422  
A;Molecule type: protein  
A;Residues: 25-48 <ROE>  
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A;Title: Mass spectrometric identification of modifications to human serum albumin treat  
A;Reference number: S36882; MUID:93384321; PMID:8373198  
A;Accession: S36882  
A;Molecule type: protein  
A;Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <FIN>  
R;Kausler, E.; Spitteller, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A;Reference number: S17599; MUID:92126241; PMID:1772598  
A;Accession: S17599  
A;Molecule type: protein  
A;Residues: 25-54; 354-357; 431-447 <KAU>  
A;Note: 49-leu was also found  
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid protea  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: A45800  
A;Molecule type: protein  
A;Residues: 166-173 <CAR>  
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tr  
A;Reference number: A03239; MUID:86242180; PMID:3087352  
A;Accession: A03239  
A;Molecule type: protein  
A;Residues: 166-173, 'L' <MOG>  
R;Galliano, M.; Minichioti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, J  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A;Title: Mutations in genetic variants of human serum albumin found in Italy.  
A;Reference number: A38255; MUID:91062352; PMID:2247440  
A;Accession: C38255  
A;Molecule type: protein  
A;Residues: 76-111 <GALL>  
A;Accession: B38255  
A;Molecule type: protein  
A;Residues: 82-105, 'K', 107-110 <GALL>  
A;Note: this variant is designated albumin Vibo Valentia  
A;Accession: A38255  
A;Molecule type: protein  
A;Residues: 76-83, 'K', 85-106 <GALL>  
A;Note: this variant is designated albumin Torino  
R;Minichioti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A;Title: The structural characterization and bilirubin-binding properties of albumin He  
A;Reference number: S33298; MUID:93292504; PMID:8513793  
A;Accession: S33298  
A;Molecule type: protein  
A;Residues: 255-263, 'E', 265-281 <MIN1>  
A;Note: this variant is designated albumin Herborn  
R;Minichioti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta, F  
Biochem. Biophys. Acta 1119, 232-238, 1992  
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe  
A;Reference number: S21078; MUID:92190239; PMID:1347703  
A;Accession: S21078  
A;Molecule type: protein  
A;Residues: 354-356, 'K', 358-378 <MIN2>  
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported  
R;He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A;Title: Atomic structure and chemistry of human serum albumin.  
A;Reference number: A46756; MUID:92334427; PMID:1630489  
A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.  
in: The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
A;Reference number: A94442  
A;Contents: annotation; three-dimensional structure and disulfide bonds  
R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A;Title: Disulfide bonds in human serum albumin.  
A;Reference number: A90930  
A;Contents: annotation; disulfide bonds  
R;Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A;Reference number: A90299; MUID:78186630; PMID:656055  
A;Contents: annotation; bilirubin-binding site  
R;Peters, T.; Reed, R.G.  
in: Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,  
A;Title: Serum albumin: conformation and active sites.  
A;Reference number: A94408

A:Contents: annotation; binding sites  
 R:Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A:Reference number: A90028; MUID:8327982; PMID:6192711  
 A:Contents: annotation; gene position  
 R:Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A:Reference number: A46755; MUID:76257808; PMID:955075  
 A:Contents: annotation  
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)  
 R:Bohney, J.P.; Fonda, M.L.; Faldutloff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate  
 A:Reference number: A56294; MUID:92183881; PMID:1544460  
 A:Contents: annotation  
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in P  
 atase activity  
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C:Comment: A large number of variants of human serum albumin have been described.  
 C:Genetics:  
 A:Gene: GDB:ALB  
 A:Cross-references: GDB:118990; OMIM:103600  
 A:Map position: 4q11-q13  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 P:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-609/Product: serum albumin #status experimental <MPT>  
 F:29-202/Domain: serum albumin repeat homology <SAI>  
 F:166-174/Product: kinetensin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:27/Binding site: copper (His) #status predicted  
 F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
 F:714/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 38; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 4.2; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0

Qy 1 AAADPHE 7  
 |||||  
 Db 386 AAADPHE 392

RESULT 3  
 D71311  
 Hypothetical protein TP0552 - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C:Accession: D71311  
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A:Reference number: A71250; MUID:9833270; PMID:9665876  
 A:Accession: D71311  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-177 <COL>  
 A:Cross-references: GB:AE001230; GB:AE000520; NID:G3322838; PIDN:AAAC65538.1; PID:G332284  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0552

Query Match 89.5%; Score 34; DB 2; Length 177;  
 Best Local Similarity 85.7%; Pred. No. 7.4; Indels 1; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 1

Qy 1 AAADPHE 7  
 |||||  
 Db 85 AAADPHE 91

RESULT 4  
 C70781  
 Probable fprB protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Mar-2000  
 C:Accession: C70781  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: C70781  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-575 <COL>  
 A:Cross-references: GB:Z73101; GB:AL123456; NID:G3261565; PIDN:CAA97393.1; PID:e2411991;  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: fprB  
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h  
 F:2-64/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 89.5%; Score 34; DB 2; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 27; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0

Qy 2 AADPHE 7  
 |||||  
 Db 270 AADPHE 275

RESULT 5  
 E69420  
 hydrogenase expression/formation regulatory protein (hypF) homolog - Archaeoglobus fulg  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 14-Apr-2003  
 C:Accession: E69420  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: E69420  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-740 <KLE>  
 A:Cross-references: GB:AE001009; GB:AE000782; NID:G2689332; PIDN:AA889876.1; PID:G26492  
 C:Superfamily: carbamoyl phosphate-converting enzyme ([Nife]-hydrogenase maturation fac

Query Match 89.5%; Score 34; DB 2; Length 740;  
 Best Local Similarity 85.7%; Pred. No. 16; Indels 0; Gaps 0;  
 Matches 6; Conservative 1; Mismatches 0

Qy 1 AAADPHE 7  
 |||||  
 Db 606 AAADPHE 612

RESULT 6  
 H5392  
 Hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: H75392  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: H75392  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-238 <WHI>  
 A:Cross-references: GB:AE001590; GB:AE000513; NID:g6459214; PIDN:AAF11027.1; PID:g6459214  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1454  
 A:Map position: 1

Query Match 86.8%; Score 33; DB 2; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPH 6  
 DB 162 AAADPH 167  
 |||||

RESULT 7  
 F86147  
 hypothetical protein TUN6.5 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: F86147  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: F86147  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-308 <STO>  
 A:Cross-references: GB:AE005172; NID:g8671835; PIDN:AAF78398.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 86.8%; Score 33; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPH 6  
 DB 117 AAADPH 122  
 |||||

RESULT 8  
 AH3101  
 IS5 family transposase tnp [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C:Accession: AH3101  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Kaip, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 star, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AH3101  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-345 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AA145230.1; PID:gl7742912; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: tnp  
 A:Map position: linear chromosome

Query Match 86.8%; Score 33; DB 2; Length 345;  
 Best Local Similarity 85.7%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAADPH 7  
 DB 167 ALADPH 173  
 |||||

RESULT 9  
 B98185  
 hypothetical protein AGR\_L\_863gl [imported] - Agrobacterium tumefaciens (strain C58, Ce  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C:Accession: B98185  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourollo, B.; Goldman  
 A.; Liu, F.; Wollam, C.; Allinger, D.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: B98185  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-351 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK89004.1; PID:gl5158792; GSPDB:GN00170  
 C:Genetics:  
 A:Gene: AGR\_L\_863gl  
 A:Map position: linear chromosome

Query Match 86.8%; Score 33; DB 2; Length 351;  
 Best Local Similarity 85.7%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAADPH 7  
 DB 173 ALADPH 179  
 |||||

RESULT 10  
 C70859  
 Probable hexosyltransferase (EC 2.4.1.1-) Rv3032 [similarity] - Mycobacterium tuberculosis  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 16-Jun-2000  
 C:Accession: C70859  
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, I.  
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: C70859  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-414 <COL>  
 A:Cross-references: GB:AL021287; GB:AL123456; NID:g3261508; PIDN:CAA16117.1; PID:g27916  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv3032  
 C:Superfamily: probable hexosyltransferase ytxN  
 C:Keywords: Glycosyltransferase; hexosyltransferase

```

Query Match      86.8%; Score 33; DB 1; Length 414;
Best Local Similarity 85.7%; Pred.No. 31;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      1      AAADPHE 7
      || || || || ||
Db      66      AAQDPHE 72

RESULT 11
E87123
Probable transferase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
C:Accession: E87123
R:Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
R.; Davies, R.W.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: E87123
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-438 <STO>
A:Cross-references: GB:AL450380; NID:g13093472; PIDN:CAC30668.1; GSPDB:GN00147
C:Genetics:
C:Superfamily: probable hexosyltransferase ytxN

```

```

Query Match      86.8%; Score 33; DB 2; Length 438;
Best Local Similarity 85.7%; Pred.No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAADPHE 7
        |||
Db       90 AAQDPHE 96

RESULT 12
A69504
X-pro aminopeptidase (pepQ) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: A69504
R.; Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlbeck, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69504
A>Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-363 <LUE>
A:Cross-references: GB:AE000962; GS:AE000782; NID:g2689285; PIDN:AAB89220.1; PID:g364850
C:Superfamily: X-pro aminopeptidase

```

```

Query Match      84.2%; Score 32; DB 2; Length 363;
Best local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
          : : : : :
Db      199 SSADPHE 205

RESULT 13
AH3200
hydroxypyruvate reductase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl

```

C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C/Accession: AH3200  
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.  
erace, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2311-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A/Reference number: AB2577; UID:21608550; PMID:11743193  
A/Accession: AH3200  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-421 <KUR>  
A/Cross-references: GB:AE008687; PIDN:AAL46022.1; PID:g17743779; GSPDB:GN00188  
A/Experimental source: strain C58 (Dupont)

C:Genetics:  
A:Gene: Atu5334  
A:Genome: plasmid  
C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1021

|                       |       |               |       |                                 |
|-----------------------|-------|---------------|-------|---------------------------------|
| Query Match           | 84.2% | Score 32;     | DB 2; | Length 421;                     |
| Best Local Similarity | 85.7% | Pred. No. 51; |       |                                 |
| Matches               | 6;    | Conservative  | 1;    | Mismatches 0; Indels 0; Gaps 0; |

Qy 1 AAADPHE 7  
|||  
|||  
Db 20 AAADPYE 26

RESULT 14

C72631

hypothetical protein APE1506 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C:Accession: C72631

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Jin-no, K.; Takai, A.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A7450; MUID:99310339; PMID:10382966

A:Accession: C72631

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <KAW>

A:Cross-references: DDBJ:AP000061; NID:95104821; PIDN:BAAB0505.1; PID:dl044291; PID:951

A:Experimental source: strain K1

C:Genetics:  
C:Gene: APE1506

```

Query Match      81.6%; Score 31; DB 2; Length 105;
Best Local Similarity 83.3%; Pred.No.18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 AADPHE 7
      |||||
Db      25 AADPHD 30

RESULT 15
S72623
ribulose-5-phosphate-epimerase (EC 5.1.-.-) - yeast (Saccharomyces cerevisiae)
N|Alternate names: POS18 protein; protein J0731; protein YUL21c
C|Species: Saccharomyces cerevisiae
C|Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Nov-1999
C|Accession: S72623; S51587; S56902
R|Juhnke, H.; Krebs, B.; Koetter, P.; Entian, K.D.
Mol. Gen. Genet. 252, 456-454, 1996
A|Title: Mutant that show increased sensitivity to hydrogen peroxide reveal an important
A|Reference number: S72623; MUID:97033550; PMID:8879247
A|Accession: S72623
A|Molecule type: DNA

```

A;Residues: 1-238 <JUH>  
 A;Cross-references: EMBL:X83571; NID:G609673; PIDN:CAA8554.1; PID:G609674  
 R;Juhnke, H.  
 submitted to the EMBL Data Library, December 1994  
 A;Reference number: S51587  
 A;Accession: S51587  
 A;Molecule type: DNA  
 A;Residues: 1-238 <JUJ>  
 A;Cross-references: EMBL:X83571; NID:G609673; PIDN:CAA8554.1; PID:G609674  
 R;Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.  
 submitted to the Protein Sequence Database, September 1995  
 A;Reference number: S56891  
 A;Accession: S56902  
 A;Molecule type: DNA  
 A;Residues: 1-238 <CZI>  
 A;Cross-references: EMBL:Z49396; NID:G1008312; PIDN:CAA89415.1; PID:G1008313; GSPDB:GN00  
 C;Genetics:  
 A;Gene: SGD:RPE1; POS18; SGD:S0003657  
 A;Cross-references: MIPS:YJL121c; SGD:S0003657  
 A;Map position: 10L  
 C;Superfamily: yeast ribulose-5-phosphate-epimerase  
 C;Keywords: carbohydrate metabolism; isomerase; leucine zipper  
  
 Query Match 81.6%; Score 31; DB 1; Length 238;  
 Best Local Similarity 83.3%; Pred. No. 44;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2 AADPHE 7  
 Db 213 AADPHD 218  
  
 Search completed: April 19, 2004, 12:02:30  
 Job time : 1.85319 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.510619 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_362\_368

Perfect score: 38

Sequence: 1 AADPHE 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description        |
|------------|-------|-------------|--------|---------------|--------------------|
| 1          | 38    | 100.0       | 600    | 1 ALBU_MACMU  | Q28522 macaca mula |
| 2          | 38    | 100.0       | 609    | 1 ALBU_HUMAN  | P02768 homo sapien |
| 3          | 34    | 89.5        | 177    | 1 Y552_TREPA  | O83563 treponema p |
| 4          | 34    | 89.5        | 575    | 1 FRAB_MYCTU  | Q10547 mycobacteri |
| 5          | 31    | 81.6        | 238    | 1 RPE_YEAST   | P46969 saccharomyc |
| 6          | 31    | 81.6        | 263    | 1 PROC_TREPA  | P27771 treponema p |
| 7          | 31    | 81.6        | 350    | 1 G3P_METKA   | P58839 methanopyru |
| 8          | 31    | 81.6        | 504    | 1 PRDB_HUMAN  | Q95q75 homo sapien |
| 9          | 31    | 81.6        | 713    | 1 ICAL_PIG    | P12675 sus scrofa  |
| 10         | 31    | 81.6        | 742    | 1 PKWA_THECU  | P49695 thermomonos |
| 11         | 31    | 81.6        | 1316   | 1 RPOC_MYCLE  | P37661 mycobacteri |
| 12         | 31    | 81.6        | 1316   | 1 RPOC_MYCTU  | P47769 mycobacteri |
| 13         | 31    | 81.6        | 1863   | 1 BRC1_HUMAN  | P38398 homo sapien |
| 14         | 31    | 81.6        | 1863   | 1 BRC1_PANTR  | Q95kx8 pan troglod |
| 15         | 30    | 78.9        | 232    | 1 Y726_MYCTU  | Q50702 mycobacteri |
| 16         | 30    | 78.9        | 393    | 1 Y762_TREPA  | O83743 treponema p |
| 17         | 30    | 78.9        | 455    | 1 TBB_DICDI   | P32256 dictyosteli |
| 18         | 30    | 78.9        | 538    | 1 ADESC_METH  | O28952 methanobact |
| 19         | 30    | 78.9        | 556    | 1 ADESC_ARCFU | O29999 archaeoglob |
| 20         | 30    | 78.9        | 569    | 1 NIR_MAIZE   | P17847 zea mays (m |
| 21         | 30    | 78.9        | 877    | 1 SYV_METH    | O26861 methanobact |
| 22         | 30    | 78.9        | 887    | 1 FTSK_RHILO  | Q98eh3 rhizobium l |
| 23         | 30    | 78.9        | 922    | 1 GVEA_AERSA  | P48369 aeromonas s |
| 24         | 30    | 78.9        | 1433   | 1 DPO3_BACUD  | Q9ka72 bacillus ha |
| 25         | 29    | 76.3        | 130    | 1 ODOI_HOUSE  | Q60597 mus musculu |
| 26         | 29    | 76.3        | 144    | 1 G10_XENLA   | P12805 xenopus lae |
| 27         | 29    | 76.3        | 146    | 1 HBB_CRONI   | P02129 crocodylus  |
| 28         | 29    | 76.3        | 194    | 1 PUR1_LACCA  | P35853 lactobacill |
| 29         | 29    | 76.3        | 209    | 1 ATI3_VACCV  | P21117 vaccinia vi |
| 30         | 29    | 76.3        | 314    | 1 SIX3_CHICK  | O42406 gallus gall |
| 31         | 29    | 76.3        | 317    | 1 UL14_HSVB   | P28949 equine herp |
| 32         | 29    | 76.3        | 322    | 1 ATI2_VACCC  | P21114 vaccinia vi |
| 33         | 29    | 76.3        | 344    | 1 HIFN_BRARE  | P59723 brachydanio |

#### ALIGNMENTS

##### RESULT 1

ALBU\_MACMU STANDARD; PRT; 600 AA.

AC Q28522;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor (fragment).

GN ALB.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

OC Cercopitheidae; Macaca.

OX NCBI\_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93211971; PubMed=8460152;

RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,

RA Daxlet J., Putnam F.W.;

RT "cDNA and protein sequence of polymorphic macaque albumins that differ

in bilirubin binding.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).

CC -I- FUNCTION: Serum albumin, the main protein of plasma, has a good

binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

hormones, bilirubin and drugs. Its main function is the regulation

of the colloidal osmotic pressure of blood.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- TISSUE SPECIFICITY: Plasma.

CC -I- SIMILARITY: Belongs to the ALB/APP/VDB family.

CC -I- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL: M90463; AAA36906.1; -

PIR: A47391; A47391.

HSP: P02768; IE7B.

InterPro: IPR000264; Serum\_albumin.

Pfam: PF00273; transport\_prot; 3.

PRINTS: PR00802; SERUMALBUMIN.

ProDom: PD002486; Serum\_albumin; 1.

SMART: SM00103; ALBUMIN; 3.

PROSITE: PS00212; ALBUMIN; 3.

Metal-binding; lipid-binding; Repeat; Signal; Copper.

NON TER 1

SIGNAL <1 10 BY SIMILARITY.

PROPEP 11 16 BY SIMILARITY.

CHAIN 17 600 SERUM ALBUMIN.

DOMAIN 17 197 ALBUMIN 1.

DOMAIN 204 389 ALBUMIN 2.

DOMAIN 396 587 ALBUMIN 3.

Q50727 m probable  
P41385 bombyx mori  
Q8p775 xanthomonas  
Q95078 drosophila  
Q35681 mus musculu  
P40748 rattus norv  
Q9bqg1 homo sapien  
P05314 spinacia ol  
Q9ph47 xyiella fas  
Q8f63 xyiella fas  
P40991 saccharomyc  
P38769 saccharomyc



|    |          |     |                         |    |   |
|----|----------|-----|-------------------------|----|---|
| FT | METAL    | 19  | COPPER (BY SIMILARITY). | RC | TISSUE=Fetal liver;   |
| FT | BINDING  | 256 | BILIRUBIN (POTENTIAL).  | RA | Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,   |
| FT | DISULFID | 69  | BY SIMILARITY.          | RA | Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;            |
| FT | DISULFID | 91  | BY SIMILARITY.          | RT | "Functional prediction of the coding sequences of 121 new genes       |
| FT | DISULFID | 106 | BY SIMILARITY.          | RT | deduced by analysis of cDNA clones from human fetal liver.;"          |
| FT | DISULFID | 140 | BY SIMILARITY.          | RL | Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 184 | BY SIMILARITY.          | RN | [6]   |
| FT | DISULFID | 184 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.                  |
| FT | DISULFID | 216 | BY SIMILARITY.          | RA | Huang M.C., Wu H.T.;  |
| FT | DISULFID | 261 | BY SIMILARITY.          | RT | "The cDNA sequences of human serum albumin.;"                         |
| FT | DISULFID | 281 | BY SIMILARITY.          | RL | Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 294 | BY SIMILARITY.          | RN | [7]   |
| FT | DISULFID | 332 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A.  |
| FT | DISULFID | 377 | BY SIMILARITY.          | RC | TISSUE=Liver, and Skeletal muscle;                                    |
| FT | DISULFID | 377 | BY SIMILARITY.          | EX | MEDLINE=22388257; PubMed=12477932;                                    |
| FT | DISULFID | 385 | BY SIMILARITY.          | RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,              |
| FT | DISULFID | 408 | BY SIMILARITY.          | RA | Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,   |
| FT | DISULFID | 453 | BY SIMILARITY.          | RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,     |
| FT | DISULFID | 477 | BY SIMILARITY.          | RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,       |
| FT | DISULFID | 493 | BY SIMILARITY.          | RA | Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,         |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Brayton M.J., Usdin T.B., Toshiyuki S., Carrinchi P., Prange C.,      |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., |
| FT | DISULFID | 530 | BY SIMILARITY.          | RA | Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,   |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,         |
| FT | DISULFID | 574 | BY SIMILARITY.          | RA | Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,              |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,        |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Generation and initial analysis of more than 15,000 full-length      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | human and mouse cDNA sequences.;"                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [8]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=76187907; PubMed=1225573;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Meloun B., Moravsek L., Kostka V.;                                    |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Complete amino acid sequence of human serum albumin.;"               |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | FEBS Lett. 58:134-137(1975).  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [9]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Brown J.R., Shockley P., Behrens P.Q.;                                |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | (in) Bing D.H. (eds.);  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | The chemistry and physiology of the human plasma proteins, pp.23-40,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Pergamon Press, New York (1979).                                      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [10]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 1-455 FROM N.A.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RC | TISSUE=Liver;   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Menaya J., Parrilla R., Ayuso M.S.;                                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [11]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 1-26 FROM N.A.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=86140099; PubMed=2419329;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Urano Y., Watanabe K., Sakai M., Tamaoki T.;                          |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "The human albumin gene. Characterization of the 5' and 3' flanking   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | regions and the polymorphic gene transcripts.;"                       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | J. Biol. Chem. 261:3244-3251(1986).                                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [12]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 222-229.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=76257808; PubMed=955075;                                      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Walker J.E.;  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Lysine residue 199 of human serum albumin is modified by             |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | acetylsalicylic acid.;"   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | FEBS Lett. 66:173-175(1976).  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [13]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-44 AND 480-499.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RC | TISSUE=Heart;   |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=95203287; PubMed=7895732;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "The human myocardial two-dimensional gel protein database: update    |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | 1994.;"   |

|    |          |     |                         |    |   |
|----|----------|-----|-------------------------|----|---|
| FT | METAL    | 19  | COPPER (BY SIMILARITY). | RC | TISSUE=Fetal liver;   |
| FT | BINDING  | 256 | BILIRUBIN (POTENTIAL).  | RA | Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,   |
| FT | DISULFID | 69  | BY SIMILARITY.          | RA | Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;            |
| FT | DISULFID | 91  | BY SIMILARITY.          | RT | "Functional prediction of the coding sequences of 121 new genes       |
| FT | DISULFID | 106 | BY SIMILARITY.          | RT | deduced by analysis of cDNA clones from human fetal liver.;"          |
| FT | DISULFID | 140 | BY SIMILARITY.          | RL | Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 184 | BY SIMILARITY.          | RN | [6]   |
| FT | DISULFID | 184 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.                  |
| FT | DISULFID | 216 | BY SIMILARITY.          | RA | Huang M.C., Wu H.T.;  |
| FT | DISULFID | 261 | BY SIMILARITY.          | RT | "The cDNA sequences of human serum albumin.;"                         |
| FT | DISULFID | 281 | BY SIMILARITY.          | RL | Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 294 | BY SIMILARITY.          | RN | [7]   |
| FT | DISULFID | 332 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A.  |
| FT | DISULFID | 377 | BY SIMILARITY.          | RC | TISSUE=Liver, and Skeletal muscle;                                    |
| FT | DISULFID | 377 | BY SIMILARITY.          | EX | MEDLINE=22388257; PubMed=12477932;                                    |
| FT | DISULFID | 385 | BY SIMILARITY.          | RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,              |
| FT | DISULFID | 408 | BY SIMILARITY.          | RA | Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,   |
| FT | DISULFID | 453 | BY SIMILARITY.          | RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,     |
| FT | DISULFID | 477 | BY SIMILARITY.          | RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,       |
| FT | DISULFID | 493 | BY SIMILARITY.          | RA | Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,         |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Brayton M.J., Usdin T.B., Toshiyuki S., Carrinchi P., Prange C.,      |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., |
| FT | DISULFID | 530 | BY SIMILARITY.          | RA | Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,   |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,         |
| FT | DISULFID | 574 | BY SIMILARITY.          | RA | Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,              |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,        |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Generation and initial analysis of more than 15,000 full-length      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | human and mouse cDNA sequences.;"                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [8]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=76187907; PubMed=1225573;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Meloun B., Moravsek L., Kostka V.;                                    |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Complete amino acid sequence of human serum albumin.;"               |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | FEBS Lett. 58:134-137(1975).  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [9]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Brown J.R., Shockley P., Behrens P.Q.;                                |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | (in) Bing D.H. (eds.);  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | The chemistry and physiology of the human plasma proteins, pp.23-40,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Pergamon Press, New York (1979).                                      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [10]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 1-455 FROM N.A.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RC | TISSUE=Liver;   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Menaya J., Parrilla R., Ayuso M.S.;                                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [11]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 1-26 FROM N.A.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=86140099; PubMed=2419329;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Urano Y., Watanabe K., Sakai M., Tamaoki T.;                          |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "The human albumin gene. Characterization of the 5' and 3' flanking   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | regions and the polymorphic gene transcripts.;"                       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | J. Biol. Chem. 261:3244-3251(1986).                                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [12]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 222-229.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=76257808; PubMed=955075;                                      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Walker J.E.;  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Lysine residue 199 of human serum albumin is modified by             |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | acetylsalicylic acid.;"   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | FEBS Lett. 66:173-175(1976).  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [13]  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-44 AND 480-499.  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RC | TISSUE=Heart;   |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=95203287; PubMed=7895732;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "The human myocardial two-dimensional gel protein database: update    |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | 1994.;"   |

|    |          |     |                         |    |   |
|----|----------|-----|-------------------------|----|---|
| FT | METAL    | 19  | COPPER (BY SIMILARITY). | RC | TISSUE=Fetal liver;   |
| FT | BINDING  | 256 | BILIRUBIN (POTENTIAL).  | RA | Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,   |
| FT | DISULFID | 69  | BY SIMILARITY.          | RA | Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;            |
| FT | DISULFID | 91  | BY SIMILARITY.          | RT | "Functional prediction of the coding sequences of 121 new genes       |
| FT | DISULFID | 106 | BY SIMILARITY.          | RT | deduced by analysis of cDNA clones from human fetal liver.;"          |
| FT | DISULFID | 140 | BY SIMILARITY.          | RL | Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 184 | BY SIMILARITY.          | RN | [6]   |
| FT | DISULFID | 184 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.                  |
| FT | DISULFID | 216 | BY SIMILARITY.          | RA | Huang M.C., Wu H.T.;  |
| FT | DISULFID | 261 | BY SIMILARITY.          | RT | "The cDNA sequences of human serum albumin.;"                         |
| FT | DISULFID | 281 | BY SIMILARITY.          | RL | Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.               |
| FT | DISULFID | 294 | BY SIMILARITY.          | RN | [7]   |
| FT | DISULFID | 332 | BY SIMILARITY.          | RP | SEQUENCE FROM N.A.  |
| FT | DISULFID | 377 | BY SIMILARITY.          | RC | TISSUE=Liver, and Skeletal muscle;                                    |
| FT | DISULFID | 377 | BY SIMILARITY.          | EX | MEDLINE=22388257; PubMed=12477932;                                    |
| FT | DISULFID | 385 | BY SIMILARITY.          | RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,              |
| FT | DISULFID | 408 | BY SIMILARITY.          | RA | Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,   |
| FT | DISULFID | 453 | BY SIMILARITY.          | RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,     |
| FT | DISULFID | 477 | BY SIMILARITY.          | RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,       |
| FT | DISULFID | 493 | BY SIMILARITY.          | RA | Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,         |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Brayton M.J., Usdin T.B., Toshiyuki S., Carrinchi P., Prange C.,      |
| FT | DISULFID | 503 | BY SIMILARITY.          | RA | Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., |
| FT | DISULFID | 530 | BY SIMILARITY.          | RA | Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,   |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., |
| FT | DISULFID | 575 | BY SIMILARITY.          | RA | Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,         |
| FT | DISULFID | 574 | BY SIMILARITY.          | RA | Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,       |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,              |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,        |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;                  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Generation and initial analysis of more than 15,000 full-length      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | human and mouse cDNA sequences.;"                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).                   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [8]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | EX | MEDLINE=76187907; PubMed=1225573;                                     |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Meloun B., Moravsek L., Kostka V.;                                    |
| FT | DISULFID | 583 | BY SIMILARITY.          | RT | "Complete amino acid sequence of human serum albumin.;"               |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | FEBS Lett. 58:134-137(1975).  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [9]   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RP | SEQUENCE OF 25-609.   |
| FT | DISULFID | 583 | BY SIMILARITY.          | RA | Brown J.R., Shockley P., Behrens P.Q.;                                |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | (in) Bing D.H. (eds.);  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | The chemistry and physiology of the human plasma proteins, pp.23-40,  |
| FT | DISULFID | 583 | BY SIMILARITY.          | RL | Pergamon Press, New York (1979).                                      |
| FT | DISULFID | 583 | BY SIMILARITY.          | RN | [10]  |

RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RX Saber M.A., Stockbauer P., Moravsek L., Meloun B.;  
RA "Disulfide bonds in human serum albumin."  
RT Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobson C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin."  
RN Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second domain of serum albumin."  
RN Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=8068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M., Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations."  
RN Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2763316;  
RA Arai K., Madison J., Huss K., Ishioke N., Satoh C., Fujita M., Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese allolalbumins."  
RN Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RP VARIANTS MANAUS; OSAKA; NAGOVA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia."  
RN Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RP DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of human serum albumin whose precursor has an aberrant signal peptidase cleavage site."  
RN Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J., Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in Italy."  
RN Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RP VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M., Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin."  
RN Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOWAGOME-3 HIS-23; KOWAGOME-2 ARG-152 AND KOWAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E., Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese."  
RN Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).

RN [24]  
RP VARIANT CASEBROOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum albumin: albumin Casebrook (494 Asp-->Asn)."  
RN Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H., Rochu D., Porta F.;  
RT "Two allolalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 38; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 2.7; Mismatches 0; Gaps 0;  
Matches 7; Conservative 0; Indels 0; Gaps 0;  
QY 1 AAADPHE 7  
DB 386 AAADPHE 392

RESULT 3  
Y552\_TREPA  
ID Y552\_TREPA STANDARD; PRT; 177 AA.  
AC O83563;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein TP0552.  
GN TP0552.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nichols;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Usterback T., McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete."  
RN Science 281:375-388(1998).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AE001230; AAC65538.1; -.  
DR PIR; D71311; D71311.  
DR TIGR; TP0552; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 177 AA; 20172 MW; 59A0F0B815D006EE CRC64;  
Query Match 89.5%; Score 34; DB 1; Length 177;  
Best Local Similarity 85.7%; Pred. No. 4.9; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Indels 0; Gaps 0;  
QY 1 AAADPHE 7  
DB 85 AAADPHE 91

RESULT 4  
ID\_FPRB MYCTU STANDARD; PRT; 575 AA.  
AC Q10547;  
DT 01-OCT-1996 (Rel. 34, Created)  
DI 01-OCT-1996 (Rel. 34, Last sequence update)  
DE Probable ferredoxin/ferredoxin-NADP reductase (EC 1.18.1.2) (FNR).  
GN FPRB OR RV0886 OR MT0909 OR MTCV31.14 OR MB0910.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1].  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala P.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh J., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
RN [2].  
RP SEQUENCE FROM N.A.  
RX SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
RA Delcher A., Utterback T., Wadman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL J. Bacteriol. 184:5479-5490(2002).  
RN [3].  
RP SEQUENCE FROM N.A.  
RX SPECIES=M.bovis; STRAIN=AF2122/97;  
RX MEDLINE=22709107; PubMed=12788972;  
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Pryor M., Duhoy S., Grondin S., Lacroix C., Mensepe C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
CC -!- CATALYTIC ACTIVITY: Reduced ferredoxin + NADP(+) = oxidized  
CC ferredoxin + NADPH.  
CC -!- COFACTOR: FAD; probably binds one or two 4Fe-4S clusters.  
CC -!- SIMILARITY: In the N-terminal, belongs to the bacterial-type  
CC ferredoxin family.  
CC -!- SIMILARITY: IN THE C-TERMINAL, TO OTHER FERREDOXIN NADP  
CC REDUCTASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 273101; CAA97393.1; -  
CC DR EMBL; AE006978; AAK45151.1; -  
CC DR EMBL; BX248336; CAD93771.1; -  
CC DR PIR; C70781; C70781.  
CC DR HSP; Q45560; 1BD6.  
CC TIGR; MT0909; -

Tuberculin; RV0886; -  
DR InterPro; IPR001450; 4Fe4S ferredoxin.  
DR InterPro; IPR000813; 7Fe ferredoxin.  
DR InterPro; IPR000759; Adnrx reductase.  
DR InterPro; IPR001327; PAD\_Pyl\_redox.  
DR Pfam; PF00037; fer4; 2.  
DR PRINTS; PR00354; 7F8SFRDOXIN.  
DR PRINTS; PR00419; ADXNDTASE.  
DR PRINTS; PR00368; FADPNR.  
DR PROSITE; PS00198; 4Fe4S FERREDOXIN; 1.  
KW Oxidoreductase; Flavoprotein; NADP; FAD; Electron transport;  
KW Iron-sulfur; 4Fe-4S; Complete proteome.  
FT DOMAIN 1 83 FERREDOXIN.  
FT DOMAIN 115 575 FERREDOXIN-NADP REDUCTASE.  
FT METAL 15 9 IRON-SULFUR 1 (BY SIMILARITY).  
FT METAL 15 15 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 19 19 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 46 46 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 49 49 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 52 52 IRON-SULFUR 1 (BY SIMILARITY).  
FT METAL 56 56 IRON-SULFUR 1 (BY SIMILARITY).  
SQ SEQUENCE 575 AA; 61337 MW; 170C58599872A810 CRC64;  
Query Match 89.5%; Score 34; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AADPHE 7  
DB 270 AADPHE 275  
RESULT 5  
ID\_RPE YEAST STANDARD; PRT; 238 AA.  
AC P46569;  
DT 01-NOV-1995 (Rel. 32, Created)  
DI 01-NOV-1995 (Rel. 32, Last sequence update)  
DI 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Ribulose-phosphate 3-epimerase (EC 5.1.3.1) (Pentose-5-phosphate 3-  
DE epimerase) (PPE) (RPE).  
GN RPE1 OR POS18 OR XYL121C OR J0731.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae;  
OX NCBI\_TaxID=4932;  
RN [1].  
RP SEQUENCE FROM N.A.  
RA Junke H.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [2].  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / FY1679;  
RX MEDLINE=97103775; PubMed=8948101;  
RA Cziepluch C., Kordes B., Pujol A., Jauniaux J.-C.;  
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
RT reveals 19 open reading frames including URA2 [5' end], TRK1, PBS2,  
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,  
RT three remnant delta elements and a Ty4 transposon."  
RL Yeast 12:1471-1474(1996).  
RN [3].  
RN CHARACTERIZATION.  
RP MEDLINE=97074231; PubMed=8929392;  
RA Miosga T., Zimmermann F.K.;  
RT "Cloning and characterization of the first two genes of the non-  
RT oxidative part of the Saccharomyces cerevisiae pentose-phosphate  
RT pathway."  
RL Curr. Genet. 30:404-409(1996).  
CC -!- CATALYTIC ACTIVITY: D-ribulose 5-phosphate = D-xylulose 5-  
CC phosphate.  
CC -!- PATHWAY: Nonoxidative branch of the pentose phosphate pathway.  
CC -!- SIMILARITY: Belongs to the ribulose-phosphate 3-epimerase family.  
CC



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; AE010355; AAM01833.1; --

DR HAMAP; MF 00559; -- 1.

DR InterPro; IPR000173; GAP\_dhhydrogenase.

DR InterPro; IPR000436; GAPDH-II\_archae.

DR Pfam; PF00044; gpdh\_2.

DR Pfam; PF02800; gpdh\_C; 1.

DR ProDom; PD007761; GAPDH-II\_archae; 1.

DR TIGRFAMs; TIGR01546; GAPDH-II\_archae; 1.

DR PROSITE; PS00071; GAPDH; 1.

KW Glycolysis; Oxidoreductase; NAD; NADP; Complete proteome.

FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE (BY

FT SIMILARITY).

SQ SEQUENCE 350 AA; 38367 MW; F9BCB72702E22E02 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 350;

Best Local Similarity 83.3%; Pred. No. 42;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7

DB 177 AADPHQ 182

-----

RESULT 8

PRDB HUMAN

ID PRDB\_HUMAN STANDARD; PRT; 504 AA.

AC Q9NQV5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE PR-domain protein 11.

GN PRDM11 OR PF08.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP Yang X.-H., Huang S.;

RT "A family of novel PR-domain (PRDM) genes as candidate tumor

RL suppressors.";

CC Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Contains 1 SET domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; AF275818; AAF87244.1; --

DR Genew; HGNC:13996; PRDM11.

DR InterPro; IPR001214; SET.

DR SMART; SM00317; SET; 1.

DR PROSITE; PS50280; SET; 1.

FT DOMAIN 142 257 SET.

SQ SEQUENCE 504 AA; 57032 MW; CF591094CCF45515 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 504;

Best Local Similarity 83.3%; Pred. No. 61;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7

DB 426 ASDPHE 431

-----

RESULT 9

ICAL\_PIG

ID ICAL\_PIG STANDARD; PRT; 713 AA.

AC P12675;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Calpain inhibitor (Calpastatin).

GN CAST.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RX MEDLINE=88241021; PubMed=2837276;

RA Takano E., Maki M., Mori H., Hatanaka M., Marti T., Titani K.,

RA Kannagi R., Marti T., Titani K., Murachi T.;

RT "Pig heart calpastatin: identification of repetitive domain

RT structures and anomalous behavior in polyacrylamide gel

RT electrophoresis.";

RL Biochemistry 27:1964-1972(1988).

RN [2]

RP SEQUENCE OF 304-554 FROM N.A.

RX MEDLINE=87054580; PubMed=3780962;

RA Takano E., Maki M., Hatanaka M., Mori H., Zenita K., Sakiyama T.,

RA Kannagi R., Marti T., Titani K., Murachi T.;

RT "Evidence for the repetitive domain structure of pig calpastatin as

RT demonstrated by cloning of complementary DNA.";

RL FEBS Lett. 208:199-202(1986).

CC -!- FUNCTION: Specific inhibition of calpain (calcium-dependent

CC cysteine protease). Plays a key role in postmortem tenderization

CC of meat and have been proposed to be involved in muscle protein

CC degradation in living tissue.

CC -!- DOMAIN: Has four inhibitory domains.

CC -!- SIMILARITY: Belongs to the calpastatin family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; M20160; AAA31012.1; --

DR EMBL; M27969; AAA31003.1; --

DR PIR; A24627; A24627.

DR PIR; A28706; A28706.

DR InterPro; IPR001259; Calpain\_inhib.

DR Pfam; PF00748; Calpain\_inhib; 4.

KW Repeat; Thiol protease inhibitor.

FT REPEAT 171 223 INHIBITORY DOMAIN 1.

FT REPEAT 307 359 INHIBITORY DOMAIN 2.

FT REPEAT 447 500 INHIBITORY DOMAIN 3.

FT REPEAT 583 636 INHIBITORY DOMAIN 4.

FT CONFLICT 328 328 L > V (IN REF. 2).

SQ SEQUENCE 713 AA; 77123 MW; ASD4E8F119CE97B5 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 713;

Best Local Similarity 85.7%; Pred. No. 89;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AADPHE 7

DB 273 AADPHE 279

-----

RESULT 10

PKWA\_THECU

ID PKWA\_THECU STANDARD; PRT; 742 AA.

P49695;  
01-FEB-1996 (Rel. 33, Created)  
01-FEB-1996 (Rel. 33, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
DE Probable serine/threonine-protein kinase pKWA (EC 2.7.1.37).  
GN PKWA OR PKW1.  
OS Thermomonospora curvata.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptosporangineae; Thermomonosporaceae; Thermomonospora.  
OX NCBI\_TaxID=2020;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM 3352;  
RX MEDLINE=96200125; PubMed=8631732;  
RA Janda L., Tichy P., Spizek J., Petrcek M.;  
RT "A deduced Thermomonospora curvata protein containing  
RL serine/threonine protein kinase and WD-repeat domains.";  
CC J. Bacteriol. 178:1487-1489(1996).  
CC -!- FUNCTION: MAY PLAY A REGULATORY ROLE DURING THE COMPLEX GROWTH  
CC CYCLE AND IN SECONDARY METABOLITE PRODUCTION.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC -!- SIMILARITY: Contains 7 WD repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF115313; AAB05922.1; -  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00400; WD40; 7.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000018; WD40; 5.  
DR SMART; SM00320; WD40; 7.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
DR PROSITE; PS00082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS00294; WD\_REPEATS\_REGION; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
WD repeat.  
FT DOMAIN 16 266 PROTEIN KINASE.  
FT NP\_BIND 22 30 ATP (BY SIMILARITY).  
FT BINDING 44 44 ATP (BY SIMILARITY).  
FT ACT\_SITE 138 138 BY SIMILARITY.  
FT DOMAIN 301 394 PRO/GLU/ASP-RICH (SPACER).  
FT REPEAT 455 496 WD 1.  
FT REPEAT 497 538 WD 2.  
FT REPEAT 539 580 WD 3.  
FT REPEAT 581 621 WD 4.  
FT REPEAT 622 663 WD 5.  
FT REPEAT 664 705 WD 6.  
FT REPEAT 706 742 WD 7.  
SQ SEQUENCE 742 AA; 78950 MW; ACL734640DB4383D CRC64;  
  
Query Match 81.6%; Score 31; DB 1; Length 742;  
Best Local Similarity 83.3%; Pred. No. 92;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 AADPHE 7  
:|||||  
DB 444 SADPHE 449

RESULT 11  
RPOC MYCLE  
ID RPOC MYCLE STANDARD; PRT; 1316 AA.  
AC P30761;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase  
DE beta' chain) (RNA polymerase beta' subunit).  
GN RPOC OR ML1890.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93188701; PubMed=8446028;  
RA Honore N.T., Berch S., Chanteau S., Doucet-Populaire F.,  
RA Eiglmeyer K., Garnier T., Georges C., Launois P., Limpitboon T.,  
RA Newton S., Niang K., del Portillo P., Ramesh G.R., Reddi P.,  
RA Ridet P.R., Sittisombut N., Wu-Hunter S., Cole S.T.;  
RT "Nucleotide sequence of the first cosmid from the Mycobacterium  
RT leprae genome project: structure and function of the Rif-Str  
RT regions.";  
RL Mol. Microbiol. 7:207-214(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltingworth T., Fraser A., Hamlin N.,  
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrell B.G.;  
RT "Massive gene decay in the leprosy bacillus.";  
RL Nature 409:1007-1011(2001).  
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
CC of DNA into RNA using the four ribonucleoside triphosphates as  
CC substrates.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC (RNA) (N).  
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core  
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1  
CC beta' chain.  
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z14314; CAA78669.1; -  
DR EMBL; AL583923; CAC30844.1; -  
DR PIR; D87145; D87145.  
DR PIR; S31146; S31146.  
DR HSP; Q9KWU6; IHQM.  
DR Leproma; ML1890; -  
DR InterPro; IPR000722; RNA\_pol\_A.  
DR InterPro; IPR007080; RNA\_pol\_Rpb1\_1.  
DR InterPro; IPR007066; RNA\_pol\_Rpb1\_3.  
DR InterPro; IPR007083; RNA\_pol\_Rpb1\_4.  
DR InterPro; IPR007081; RNA\_pol\_Rpb1\_5.  
DR InterPro; IPR006592; RNA\_pol\_N.  
DR Pfam; PF04997; RNA\_pol\_Rpb1\_1; 1.  
DR Pfam; PF00623; RNA\_pol\_Rpb1\_2; 1.  
DR Pfam; PF04983; RNA\_pol\_Rpb1\_3; 1.

DR Pfam; PF05000; RNA\_pol\_Rpb1\_4; 1.  
DR Pfam; PF04998; RNA\_pol\_Rpb1\_5; 1.  
DR SMART; SM00663; RPOA\_N; 1.  
KW Transferase; DNA-directed RNA polymerase; Transcription;  
KW Complete proteome.  
FT CONFLICT 598 A -> R (IN REF. 1).  
SQ SEQUENCE 1316 AA; 146895 MW; 7131AB63B53F6BCF CRC64;  
  
Query Match 81.6%; Score 31; DB 1; Length 1316;  
Best Local Similarity 83.3%; Pred No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 AADPHE 7  
Db 1115 SADPHE 1120  
:|||||  
-----  
RESULT 12  
ID RPOC MYCTU STANDARD; PRT; 1316 AA.  
AC P47769; O06771;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 15-FEB-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase  
beta' chain) (RNA polymerase beta' subunit).  
GN RPOC OR RV0668 OR MT0696 OR MTC1376.07C OR MB0687.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=3634230;  
RA Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekala F.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Connor R.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagsis K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL J. Bacteriol. 184:5479-5490(2002).  
RN [3]  
RP SEQUENCE OF 1-148 FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37RV;  
RX MEDLINE=94304130; PubMed=8031050;  
RA Miller L.P., Crawford J.T., Shinnick T.M.;  
RT "The rpoB gene of Mycobacterium tuberculosis."  
RL Antimicrob. Agents Chemother. 38:805-811(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.bovis; STRAIN=AF122/97;  
RX MEDLINE=22709107; PubMed=12789972;  
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
of DNA into RNA using the four ribonucleoside triphosphates as  
substrates.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
{RNA}(N).  
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core  
enzyme which is composed of 2 alpha chains, 1 beta chain, and 1  
beta' chain.  
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; Z59572; CAB09389.1; --  
EMBL; AE006964; AAK44922.1; --  
EMBL; L27989; AAK21417.1; --  
EMBL; BX248336; CAD93549.1; --  
PIR; G70535; G70535.  
HSSP; Q9KWD6; IHQM.  
TIGR; MT0696; --  
DR TubercuList; RV0668; --  
DR InterPro; IPR007072; RNA\_pol\_A.  
DR InterPro; IPR007080; RNA\_pol\_Rpb1\_1.  
DR InterPro; IPR007066; RNA\_pol\_Rpb1\_3.  
DR InterPro; IPR007083; RNA\_pol\_Rpb1\_4.  
DR InterPro; IPR007081; RNA\_pol\_Rpb1\_5.  
DR InterPro; IPR006592; RNA\_pol\_A\_N.  
DR Pfam; PF04997; RNA\_pol\_Rpb1\_1; 1.  
DR Pfam; PF00623; RNA\_pol\_Rpb1\_2; 1.  
DR Pfam; PF04983; RNA\_pol\_Rpb1\_3; 1.  
DR Pfam; PF05000; RNA\_pol\_Rpb1\_4; 1.  
DR Pfam; PF04998; RNA\_pol\_Rpb1\_5; 1.  
DR SMART; SM00663; RPOA\_N.1  
KW Transferase; DNA-directed RNA polymerase; Transcription;  
KW Complete proteome.  
FT CONFLICT 594 594 G -> E (IN REF. 2).  
SQ SEQUENCE 1316 AA; 146769 MW; 45BF24839AF53E8B CRC64;  
  
Query Match 81.6%; Score 31; DB 1; Length 1316;  
Best Local Similarity 83.3%; Pred No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 AADPHE 7  
Db 1115 SADPHE 1120  
:|||||  
-----  
RESULT 13  
ID BRCL HUMAN STANDARD; PRT; 1863 AA.  
AC P38398;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Breast cancer type 1 susceptibility protein.  
GN BRCL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT ARG-1775.  
RX MEDLINE=95025996; PubMed=7545954;  
RA Miki Y., Swensen J., Shattuck-Bidens D., Futreal P.A., Harshman K.,  
RA Tavtigian S., Liu Q., Cochran C., Bennett L.M., Ding W., Bell R.,

RA Rosenthal J., Husey C., Tran T., McClure M., Frye C., Hattier T.,  
RA Shelp R., Haugen-Strano A., Kacher H., Yakumo K., Gholami Z.,  
RA Shaffer D., Stone S., Bayer S., Wray C., Bogden R., Dayananth P.,  
RA Ward J., Tonin P., Narod S., Bristow P.K., Norris F.H., Helvaring L.,  
RA Morrison P., Rostek P., Lai M., Barrett J.C., Lewis C., Neuhausen S.,  
RA Cannon-Albright L., Goddard D., Wiseman R., Kamb A., Skolnick M.H.,  
RT "A strong candidate for the breast and ovarian cancer susceptibility  
RT gene BRCA1".  
RL Science 266:66-71(1994).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=97092885; PubMed=8939427;  
RX Smith T.M., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,  
RA Hood L., King M.-C.;  
RT "Complete genomic sequence and analysis of 117 kb of human DNA  
RT containing the gene BRCA1".  
RL Genome Res. 6:1029-1049(1996).  
RN [3]  
RN FUNCTION AS A E2-DEPENDENT UBIQUITIN LIGASE.  
RP MEDLINE=99432238; PubMed=10500182;  
RX Loric K.L., Jensen J.P., Fang S., Ong A.M., Hatakeyama S.,  
RA Weissman A.M.;  
RT "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent  
RT ubiquitination".  
RL Proc. Natl. Acad. Sci. U.S.A. 96:11364-11369(1999).  
RN [4]  
RN IDENTIFICATION OF BRCA1 AS MEMBER OF BASC.  
RP MEDLINE=20245492; PubMed=10783165;  
RX Wang Y., Cortez D., Yazdi P., Neff N., Ellledge S.J., Qin J.;  
RA "BASC, a super complex of BRCA1-associated proteins involved in the  
RT recognition and repair of aberrant DNA structures".  
RL Genes Dev. 14:927-939(2000).  
RN [5]  
RN INTERACTION WITH SMC1L1.  
RP MEDLINE=21866464; PubMed=11877377;  
RX Yazdi P.T., Wang Y., Zhao S., Patel N., Lee E.Y.-H.P., Qin J.;  
RA "SMC1 is a downstream effector in the ATM/ATR1 branch of the human  
RT S-phase checkpoint".  
RL Genes Dev. 15:571-582(2002).  
RN [6]  
RN INTERACTION WITH NLF3.  
RP MEDLINE=21601803; PubMed=11739404;  
RX Ye Q., Hu Y.-F., Zhong H., Nye A.C., Belmont A.S., Li R.;  
RA "BRCA1-induced large-scale chromatin unfolding and allele-specific  
RT effects of cancer-predisposing mutations".  
RL J. Cell Biol. 155:911-921(2001).  
RN [7]  
RN PHOSPHORYLATION BY ATM, AND MUTAGENESIS OF SER-1387; SER-1423 AND  
RP SER-1524.  
RX Xu B., O'Donnell A.H., Kim S.-T., Kastan M.B.;  
RA "Phosphorylation of serine 1387 in BRCA1 is specifically required for  
RT the Atm-mediated S-phase checkpoint after ionizing irradiation".  
RL Cancer Res. 62:4588-4591(2002).  
RN [8]  
RN REVIEW ON VARIANTS.  
RP MEDLINE=96400954; PubMed=8807330;  
RX Couch F.J., Weber B.L.;  
RA "Mutations and polymorphisms in the familial early-onset breast  
RT cancer (BRCA1) gene".  
RL Hum. Mutat. 8:8-18(1996).  
RN [9]  
RN VARIANTS LEU-1637; GLU-1708 AND ARG-1775.  
RX MEDLINE=95025878; PubMed=7939630;  
RA Futreal P.A., Liu Q., Shattuck-Eidens D., Cochran C., Harshman K.,  
RA Tavtigian S., Bennett L.M., Haugen-Strano A., Swensen J., Miki Y.,  
RA Edgington K., McClure M., Frye C., Weaver-Fellhaus J., Ding W.,  
RA Ghosami Z., Soedarkrist P., Terry L., Jhanwar S., Berchuk A.,  
RA Iglehart J.D., Marks J., Ballinger D.G., Barrett J.C., Skolnick M.H.,  
RA Kamb A., Wiseman R.;  
RT "BRCA1 mutations in primary breast and ovarian carcinomas".  
RL Science 266:120-122(1994).  
RN [10]  
RP VARIANT BC GLY-64, AND VARIANTS ALA-772; ASN-1040 AND GLY-1443.  
RX MEDLINE=95201806; PubMed=7894491;  
RA Castilla L.H., Couch F.J., Erdos M.R., Hoskins K.F., Calzone K.,  
RA Garber J.E., Boyd J., Lubin M.B., Deshano M.L., Brody L.C.,  
RA Collins F.S., Weber B.L.;  
RT "Mutations in the BRCA1 gene in families with early-onset breast and  
RT ovarian cancer".  
RL Nat. Genet. 8:387-391(1994).  
RN [11]  
RN VARIANT BC GLY-61, AND VARIANTS ARG-356; GLY-1038; ASN-1040; ARG-1183  
RP AND GLY-1613.  
RX MEDLINE=95201808; PubMed=7894493;  
RA Friedman L.S., Ostermeyer E.A., Szabo C.I., Dowd P., Lynch E.D.,  
RA Rowell S.E., King M.-C.;  
RT "Confirmation of BRCA1 by analysis of germline mutations linked to  
RT breast and ovarian cancer in ten families".  
RL Nat. Genet. 8:399-404(1994).  
RN [12]  
RN VARIANT BC GLY-61.  
RX MEDLINE=96108965; PubMed=8554067;  
RA Serova O., Montagna M., Torchard D., Narod S.A., Tonin P., Sylla B.,  
RA Lynch H.T., Feunteun J., Lenoir G.M.;  
RT "A high incidence of BRCA1 mutations in 20 breast-ovarian cancer  
RT families".  
RL Am. J. Hum. Genet. 58:42-51(1996).  
RN [13]  
RN VARIANT BOC TRP-841.  
RX MEDLINE=97123469; PubMed=8968716;  
RA Barker D.F., Almeida E.F.A., Casey G., Paine P.R., Liao S.-Y.,  
RA Masunaka I., Noble B., Kurosaki T., Anton-Culver H.;  
RT "BRCA1 R841W: a strong candidate for a common mutation with moderate  
RT phenotype".  
RL Genet. Epidemiol. 13:595-604(1996).  
RN [14]  
RN VARIANTS BC AND BOC.  
RX MEDLINE=96372821; PubMed=8776500;  
RA Durocher F., Shattuck-Eidens D., McClure M., Labrie F.,  
RA Skolnick M.H., Goldgar D.E., Simard J.;  
RT "Comparison of BRCA1 polymorphisms, rare sequence variants and/or  
RT missense mutations in unaffected and breast/ovarian cancer  
RT populations".  
RL Hum. Mol. Genet. 5:835-842(1996).  
RN [15]  
RN VARIANTS BC MET-271 AND SER-1150.  
RX MEDLINE=96303704; PubMed=8723683;  
RA Katagiri T., Eto M., Ito I., Kobayashi K., Yoshimoto M., Iwase T.,  
RA Kasumi F., Miki Y., Skolnick M.H., Nakamura Y.;  
RT "Mutations in the BRCA1 gene in Japanese breast cancer patients".  
RL Hum. Mutat. 7:334-339(1996).  
RN [16]  
RN VARIANT BC GLY-61, AND VARIANTS ARG-239; TRP-841 AND ILE-1512.  
RX MEDLINE=98430998; PubMed=9760198;  
RA Dong J., Chang-Claude J., Wu Y., Schumacher V., Debatin I., Tonin P.,  
RA Royer-Pokora B.;  
RT "A high proportion of mutations in the BRCA1 gene in German  
RT breast/ovarian cancer families with clustering of mutations in the 3'  
RT third of the gene".  
RL Hum. Genet. 103:154-161(1998).  
RN [17]  
RN VARIANT BC GLY-64, AND VARIANTS ALA-772; GLU-820; ASN-1040; GLY-1443;  
RP ILE-1512; LEU-1637 AND ILE-1652.  
RX MEDLINE=98141685; PubMed=9482581;  
RA Andersen T.I., Eiken H.G., Couch F., Kaada G., Skrede M., Johnsen H.,  
RA Aloysius T.A., Tveit K.W., Tranbjærg L., Doerum A., Møller P.,  
RA Weber B.L., Boerresen-Dale A.-L.;  
RT "Constant denaturation gel electrophoresis (CDGE) in BRCA1 mutation  
RT screening".  
RL Hum. Mutat. 11:166-174(1998).  
RN [18]  
RN VARIANTS BC SER-22; LEU-461; ASP-465; VAL-552; SER-892; ASP-960;  
RP ILE-1025 AND ALA-1047.  
RX MEDLINE=9827917; PubMed=9609997;  
RA Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Asaishi K., Abe R.,



RA Tsuchiya A., Sugano M., Takai S., Yoneda M., Fukutomi T., Namba K.,  
RA Makita M., Okazaki H., Hirata K., Okazaki M., Furutsuma Y.,  
RA Horiguchi Y., Iino Y., Karino T., Ayabe H., Hara S., Kajiwara T.,  
RA Moriguchi S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunitomo K.,  
RA Sano H., Kurebayashi J.-I., Shimotsuma K., Nakamura Y., Waki Y.,  
RT "High proportion of missense mutations of the BRCA1 and BRCA2 genes in  
RL Japanese breast cancer families.",  
RL J. Hum. Genet. 43:42-48(1998).  
RN [19]  
RP VARIANT OC ARG-1749.  
RX MEDLINE=20455732; PubMed=10486320;  
RA Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,  
RA Ponder B.A.J.,  
RT "The contribution of germline BRCA1 and BRCA2 mutations to familial  
RT ovarian cancer: no evidence for other ovarian cancer-susceptibility  
RT genes.",  
RL Am. J. Hum. Genet. 65:1021-1029(1999).  
RN [20]  
RP VARIANT BC SER-346, AND VARIANTS LEU-871; GLY-1038; ARG-1183 AND  
RP GLY-1613.  
RX MEDLINE=99254821; PubMed=10323242;  
RA Li S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,  
RA Huang H.-W., Chen L.-M., Kao H.-W., Chen J.H., Tseng J.-N., Chen A.,  
RA Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.,  
RT "Molecular characterization of germline mutations in the BRCA1 and  
RT BRCA2 genes from breast cancer families in Taiwan.",  
RL Hum. Genet. 104:201-204(1999).  
RN [21]  
RP VARIANTS OC, AND VARIANTS.  
RX MEDLINE=99214030; PubMed=10196379;  
RA Janzic S.A., Ziogas A., Krutroy L.M., Kraemer M., Plummer S.J.,  
RA Cohen P., Gildea M., Barker D., Haile R., Casey G., Anton-Culver H.,  
RT "Germline BRCA1 alterations in a population-based series of ovarian  
RT cancer cases",  
RL Hum. Mol. Genet. 8:889-897(1999).  
CC -!- FUNCTION: Plays a central role in DNA repair by facilitating  
CC cellular response to DNA repair. Required for appropriate cell  
CC cycle arrests after ionizing irradiation in both the S-phase and

Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: Plays a central role in DNA repair by facilitating  
cellular response to DNA repair. Required for appropriate cell  
cycle arrests after ionizing irradiation in both the S-phase and  
the G2 phase of the cell cycle. Involved in transcriptional  
regulation of P1 in response to DNA damage. May function as a  
transcriptional regulator. Mediates E2-dependent ubiquitination (By  
similarity).  
-!- SUBUNIT: Part of the BRCA1-associated genome surveillance complex  
(BASC), which contains BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 and  
the RAD50-MRE11-NBS1 protein complex. This association could be a  
dynamic process changing throughout the cell cycle and within  
subnuclear domains. CtIP interacts specifically with the BRCT  
domains. Interacts with RNA polymerase II holoenzyme. Interacts  
with SMC1L1 and COBRAL/NELPB (By similarity).  
-!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
-!- PTM: Phosphorylated by ATM upon ionizing radiation (By  
similarity).  
-!- SIMILARITY: Contains 1 RING-type zinc finger.  
-!- SIMILARITY: Contains 2 BRCT domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF207822; AAG3492.1; -;  
EMBL; AY365046; AAR04849.1; -;  
GO; GO:0008274; C:gamma-tubulin ring complex; NAS.  
GO; GO:0005634; C:nucleus; TAS.  
GO; GO:0016329; F:apoptosis regulator activity; TAS.  
GO; GO:0005515; F:protein binding; IPT.  
GO; GO:0018563; F:transcriptional activator activity; TAS.  
GO; GO:0015631; F:tubulin binding; NAS.  
GO; GO:0006978; P:DNA damage response; signal transduction by . . .; TAS.  
GO; GO:0046600; P:negative regulation of centricole replication; NAS.  
GO; GO:0045739; P:positive regulation of DNA repair; NAS.  
GO; GO:0042127; P:regulation of cell proliferation; TAS.  
GO; GO:0006359; P:regulation of transcription from Pol III pr. . .; TAS.  
InterPro; IPR001357; BRCT.  
InterPro; IPR002378; Brst\_cancerI.  
InterPro; IPR001841; Znf\_ring.  
Pfam; PF00533; BRCT; 2.  
Pfam; PF00097; zf-C3HC4; 1.  
PRINTS; PR00493; BRSTCANCER1.  
SMART; SM00184; RING; 1.  
SMART; SM00292; BRCT; 2.  
PROSITE; PS00172; BRCT; 2.  
PROSITE; PS00518; ZF\_RING\_1; 1.  
PROSITE; PS00089; ZF\_RING\_2; 1.  
DNA repair; Zinc-finger; DNA-binding; Nuclear protein; Anti-oncogene;  
Repeat; Phosphorylation; Polymorphism.  
ZNF\_RING 24 65 RING-TYPE.  
DOMAIN 1642 1736 BRCT 1.  
DOMAIN 1756 1855 BRCT 2.  
DOMAIN 651 654 POLY-LYS.  
FT MOD RES 1387 1387 PHOSPHORYLATION (BY ATM) (BY SIMILARITY).  
FT MOD RES 1423 1423 PHOSPHORYLATION (BY ATM) (BY SIMILARITY).  
FT MOD\_RES 1524 1524 PHOSPHORYLATION (BY ATM) (BY SIMILARITY).  
FT VARIANT 309 309 K -> E.  
FT VARIANT 590 590 S -> G.  
FT VARIANT 731 731 K -> E.  
FT VARIANT 1100 1100 G -> E.  
FT CONFLICT 427 427 E -> K (IN REF. 2).  
FT CONFLICT 925 925 I -> T (IN REF. 2).  
FT CONFLICT 1520 1520 R -> T (IN REF. 1).  
SQ SEQUENCE 1863 AA; 207898 MW; 49673829CCFA756E CRC64;

Query Match 81.6%; Score 31; DB 1; Length 1863;

QY 2 AADPHE 7  
| : ||||  
Db 433 ASDPHE 438

```

RES001.15
YY226 MYCTU
STANDARD;          PRT;    232 AA.
Q50702;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical PPE-family protein RV3426.
RV3426 OR MTCY78.03C.
Mycobacterium tuberculosis.
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_taxid=1773;
[1]
SEQUENCE FROM N.A.
STRAIN=H37Rv;
MEDLINE=98295987; PubMed=9634230;
Gordon S.V., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Badcock K., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
Davies R., Basham D., Brown D., Chillingworth T., Connor R.,
Hornsby T., Jagels K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Oliver S., Osborne K., Krogh A., McLean J., Moule S., Murphy L.,
Rutter S., Seeger K., Quail M.A., Rajandream M.A., Rogers J.,
Rutten J.E., Taylor K., Skelton S., Squares S., Squares R.,
Sutton J.E., Whitehead S., Barrall B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
Nature 393:537-544(1998).

```

EMBL; Z77165; CAB01030.1; --  
 PIR; G70738; G70738.  
 TubercuList; RV3426; --  
 InterPro; IPR000030; Microbac\_PPE.  
 Pfam; PF00823; PPE; 1.  
 Hypothetical protein; Complete proteome.  
 SEQUENCE 232 AA; 25872 MW; D76512D49B272C6 CRC64;  
 CC

Qy 1 AADPHE 7  
| | | | |  
Db 169 AADPHE 175

Search completed: April 19, 2004, 11:52:54  
Job time : 2.51062 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.70175 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_362\_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID     | Description           |
|------------|-------|-------------|--------|--------|-----------------------|
| 1          | 38    | 100.0       | 396    | Q81UK7 | Q81uk7 homo sapien    |
| 2          | 38    | 100.0       | 417    | Q86YGO | Q86ygo homo sapien    |
| 3          | 38    | 100.0       | 4177   | Q9GUP2 | Q9gup2 caenorhabdi    |
| 4          | 34    | 89.5        | 157    | Q8NF55 | Q8nf55 homo sapien    |
| 5          | 34    | 89.5        | 310    | Q7Z3Y1 | Q7z3y1 homo sapien    |
| 6          | 34    | 89.5        | 323    | Q68908 | Q68908 streptomyc     |
| 7          | 34    | 89.5        | 323    | Q8X200 | Q8x200 ralstonia s    |
| 8          | 34    | 89.5        | 691    | Q7ZUR6 | Q7zue6 brachydanio    |
| 9          | 34    | 89.5        | 740    | Q28905 | Q28905 archaeoglob    |
| 10         | 33    | 86.8        | 180    | Q8PRL4 | Q8prl4 xanthomonas    |
| 11         | 33    | 86.8        | 238    | Q8RUD4 | Q8rud4 deinothococcus |
| 12         | 33    | 86.8        | 283    | Q8PC99 | Q8pc99 xanthomonas    |
| 13         | 33    | 86.8        | 308    | Q9LQ92 | Q9lq92 arabidopsis    |
| 14         | 33    | 86.8        | 351    | Q8U7L3 | Q8u7l3 agrobacteri    |
| 15         | 33    | 86.8        | 414    | Q53279 | Q53279 mycobacteri    |
| 16         | 33    | 86.8        | 414    | Q7TXE8 | Q7txe8 mycobacteri    |

|    |    |      |      |    |        |
|----|----|------|------|----|--------|
| 17 | 33 | 86.8 | 438  | 16 | Q9CBR3 |
| 18 | 33 | 86.8 | 459  | 10 | Q8WZB7 |
| 19 | 33 | 86.8 | 464  | 10 | Q8WZB6 |
| 20 | 33 | 86.8 | 630  | 16 | Q89X41 |
| 21 | 33 | 86.8 | 632  | 2  | Q9KW53 |
| 22 | 33 | 86.8 | 667  | 2  | Q9WXA5 |
| 23 | 33 | 86.8 | 749  | 11 | Q8C3X9 |
| 24 | 33 | 86.8 | 768  | 5  | Q97215 |
| 25 | 33 | 86.8 | 1205 | 11 | Q8CGW1 |
| 26 | 33 | 86.8 | 1574 | 11 | Q9QZC2 |
| 27 | 33 | 86.8 | 1627 | 10 | Q8LQV0 |
| 28 | 33 | 86.8 | 1939 | 2  | Q8GME1 |
| 29 | 32 | 84.2 | 130  | 2  | Q9AES9 |
| 30 | 32 | 84.2 | 130  | 2  | Q8GLR0 |
| 31 | 32 | 84.2 | 359  | 16 | Q8R642 |
| 32 | 32 | 84.2 | 363  | 17 | Q28245 |
| 33 | 32 | 84.2 | 421  | 16 | Q8U0Z0 |
| 34 | 31 | 81.6 | 101  | 16 | Q8DK11 |
| 35 | 31 | 81.6 | 105  | 17 | Q9YBU3 |
| 36 | 31 | 81.6 | 248  | 16 | Q88794 |
| 37 | 31 | 81.6 | 272  | 5  | Q45091 |
| 38 | 31 | 81.6 | 325  | 16 | Q82BG7 |
| 39 | 31 | 81.6 | 345  | 5  | P91437 |
| 40 | 31 | 81.6 | 390  | 13 | Q42223 |
| 41 | 31 | 81.6 | 418  | 10 | Q9SGS7 |
| 42 | 31 | 81.6 | 477  | 4  | Q8N9F1 |
| 43 | 31 | 81.6 | 493  | 16 | Q7V4S5 |
| 44 | 31 | 81.6 | 499  | 10 | Q9LG10 |
| 45 | 31 | 81.6 | 630  | 13 | Q57336 |

## ALIGNMENTS

### RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7; (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Similar to serum albumin precursor.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Strausberg R;

RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC035969; AAH35969.1; -

DR GO; GO:0005615; C:extracellular space; IEA.

DR GO; GO:0005386; F:carrier activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; transport P1ot; 2.

DR PRINTS; PR00802; SERUMALBUMIN.

DR SMART; SM00103; ALBUMIN; 2.

DR PROSITE; PS00212; ALBUMIN; 2.

SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 38; DB 4; Length 396;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7

Db 173 AAADPHE 179

### RESULT 2

Q86YGO

ID Q86YGO PRELIMINARY; PRT; 417 AA.  
AC Q86YGO;  
DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Similar to alpha-fetoprotein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041789; AAH41789.1; -  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SMO0103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 417 AA; 47360 MW; 16E76483BEF4E8D CRC64;  
Query Match 100.0%; Score 38; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAADPHE 7  
DB 194 AAADPHE 200  
|||||  
RESULT 3  
Q86YGO PRELIMINARY; PRT; 417 AA.  
AC Q86YGO;  
DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Hypothetical protein Y67D8C.5.  
GN Y67D8C.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA Waterston R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Edwards J., Du H., Lamar B., Kemp K., Wohldmann P., Walker C.;  
RT "The sequence of C. elegans cosmid Y67D8C.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC025724; AG23375.2; -  
DR WormPep; Y67D8C.5; CB3165.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR008938; ARM.  
DR InterPro; IPR000569; HECT domain.  
DR InterPro; IPR000449; UBA\_domain.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00627; UBA; 1.  
DR SMART; SMO0119; HECTC; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS0030; UBA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 4177 AA; 465401 MW; 9C32BF90AB87FD58 CRC64;  
Query Match 100.0%; Score 38; DB 5; Length 4177;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAADPHE 7  
DB 1431 AAADPHE 1437  
|||||  
RESULT 4  
Q8NF55 PRELIMINARY; PRT; 157 AA.  
AC Q8NF55;  
DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE FLJ00335 protein (Fragment).  
GN FLJ00335.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human  
spleen.";  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK090424; BAC03405.1; -  
DR InterPro; IPR000949; ELM2.  
DR Pfam; PF01448; ELM2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 157 AA; 17936 MW; 8DFE8C792B17D1BE CRC64;  
Query Match 89.5%; Score 34; DB 4; Length 157;  
Best Local Similarity 85.7%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAADPHE 7  
DB 32 AAADPHK 38  
|||||  
RESULT 5  
Q723Y1 PRELIMINARY; PRT; 310 AA.  
ID Q723Y1  
AC Q723Y1;  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peripheral Nervous System;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diakris L.L., Maruina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquevano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2].  
SEQUENCE FROM N.A.  
TISSUE=Peripheral Nervous System;  
Strausberg K.;  
Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BC052976; AAH52976.1; -.  
Hypothetical protein.  
SEQUENCE 310 AA; 35335 MW; 0B07FDE180821E44 CRC64;  
Query Match 89.5%; Score 34; DB 4; Length 310;  
Best Local Similarity 85.7%; Pred. No. 70;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps  
QY 1 AAADPHE 7  
|||||:  
DB 7 AAADPHK 13  
RESULT 6  
C68908 PRELIMINARY; PRT; 323 AA.  
ID C68908  
OC C68908  
DT 01-AUG-1998 (trEMBLrel. 07, Created)  
DT 01-AUG-1998 (trEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (trEMBLrel. 24, Last annotation update)  
DE Membrane protein of putative ABC transporter.  
GN FENB.  
DE Streptomyces roseofulvus.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94237486; PubMed=8181754;  
RA Bibb M.J., Sherman D.H., Omura S., Hopwood D.A.;  
RT "Cloning, sequencing and deduced functions of a cluster of  
RT Streptomyces genes probably encoding biosynthesis of the polyketide  
RT antibiotic frenolicin.";  
RL Gene 142:131-39 (1994).  
[2]  
SEQUENCE FROM N.A.  
RA Reeves C.D., Soliday C.L.;  
RT "Analysis of a 27 kb region of Streptomyces roseofulvus containing  
RT genes for frenolicin biosynthesis";  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF058302; AAC18097.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD transp.  
DR Pfam; PF00528; BPD transp. 1.  
DR SEQUENCE 323 AA; 33492 MW; 4F122511955658C8 CRC64;  
Query Match 89.5%; Score 34; DB 2; Length 323;  
Best Local Similarity 85.7%; Pred. No. 73;

```

QY 2 AADPHE 7
   |||||
Db 560 AADPHE 565

RESULT 9
O28905 PRELIMINARY; PRT; 740 AA.
AC O28905;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hydrogenase expression/formation regulatory protein (HYPF).
GN A1366.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL; AE001009; AAB89876.1; -.
DR TIGR; AF1366; -.
DR GO; GO:0003988; F:acylphosphatase activity; IEA.
DR InterPro; IPR001792; Acylphosphatase.
DR InterPro; IPR004421; HyPF.
DR InterPro; IPR006071; SUA5/ycio/yrdC_N.
DR InterPro; IPR006070; SUA5/ycio/yrdC_N.
DR Pfam; PF00708; Acylphosphatase; 1.
DR Pfam; PF01300; SUA5_ycio_yrdC; 1.
DR ProDom; PD001884; Acylphosphatase; 1.
DR ProDom; PD002209; SUA5/ycio/yrdC; 1.
DR TIGRFAMs; TIGR00143; hyPF; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
DR PIRSF; PIRSF006256; HyPF; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 740 AA; 83060 MW; 9C0D53202A64B157 CRC64;

Query Match 89.5%; Score 34; DB 17; Length 740;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADPHE 7
   |||||
Db 606 AADPHE 612

RESULT 10
Q8PRL4 PRELIMINARY; PRT; 180 AA.
AC Q8PRL4;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hypothetical protein XACB0006.
GN XACB0006.
OS Xanthomonas axonopodis (pv. citri).

Plasmid pXAC64.
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferrer J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferrero M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katayama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locals E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.H.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C. de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE008925; RAM39252.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 180 AA; 20041 MW; C317CAEEA3C31E88 CRC64;

Query Match 86.8%; Score 33; DB 16; Length 180;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AADPHE 7
   |||||
Db 102 AADPHE 108

RESULT 11
Q9RUD4 PRELIMINARY; PRT; 238 AA.
AC Q9RUD4;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein DR1454.
GN DR1454.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AE001990; AAP11027.1; -.
DR FTR; H75392; H75392.
DR TIGR; DR1454; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 238 AA; 25113 MW; 489B8C792C11E7AD CRC64;

```

```

Query Match      86.8%; Score 33; DB 16; Length 238;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
Db      162 AAADPH 167

RESULT 12
Q8PC99 PRELIMINARY; PRT; 283 AA.
AC Q8PC99;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Proteorhizobium oxidase.
GN HEMK OR XCC0835
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
CX NCBI_TaxID=340;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Ciccarelli R.M., Cannavan F., Cardoso J., Chamberg F., Chapina L.P.,
RA Ciccarelli R.M., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Maidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012183; AAM40150.1.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008276; F:protein methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR GO; GO:0006479; P:protein amino acid methylation; IEA.
DR InterPro; IPR004556; HenK.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR TIGRFAMs; TIGR000336; henKfam; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 30275 MW; A5F4E951E8E33479 CRC64;

Query Match      86.8%; Score 33; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
Db      191 AAADPH 196

RESULT 13
Q9LQ92 PRELIMINARY; PRT; 308 AA.
AC Q9LQ92;
ID Q9LQ92;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Ciccarelli R.M., Cannavan F., Cardoso J., Chamberg F., Chapina L.P.,
RA Ciccarelli R.M., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Maidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012183; AAM40150.1.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008276; F:protein methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR GO; GO:0006479; P:protein amino acid methylation; IEA.
DR InterPro; IPR004556; HenK.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR TIGRFAMs; TIGR000336; henKfam; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 30275 MW; A5F4E951E8E33479 CRC64;

Query Match      86.8%; Score 33; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
Db      191 AAADPH 196

RESULT 13
Q9LQ92 PRELIMINARY; PRT; 308 AA.
AC Q9LQ92;
ID Q9LQ92;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Ciccarelli R.M., Cannavan F., Cardoso J., Chamberg F., Chapina L.P.,
RA Ciccarelli R.M., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Maidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012183; AAM40150.1.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008276; F:protein methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR GO; GO:0006479; P:protein amino acid methylation; IEA.
DR InterPro; IPR004556; HenK.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR TIGRFAMs; TIGR000336; henKfam; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 30275 MW; A5F4E951E8E33479 CRC64;

Query Match      86.8%; Score 33; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
Db      117 AAADPH 122

RESULT 14
Q8U7L3 PRELIMINARY; PRT; 351 AA.
AC Q8U7L3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1S5 family transposase.
GN TNP OR ATU4436 OR AGR_L_863GL.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
SEQUENCE FROM N.A.
RC MEDLINE=21608550; PubMed=11743193;
RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L.,
RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323 (2001).
RN [2]
SEQUENCE FROM N.A.
RC MEDLINE=21608551; PubMed=11743194;
RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Hummel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA William C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

```

Search completed: April 19, 2004, 12:00:11  
Job time : 4.70175 secs

QY 1 AAADPHE 7  
D<sub>B</sub> 173 ALADPHE 179

SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and  
RT laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/Genbank/DBSJ databases

|    |              |           |                                   |                   |
|----|--------------|-----------|-----------------------------------|-------------------|
| DR | EMBL         | AL021287  | CAA16117.1                        | -                 |
| DR | EMBL         | AE007130  | AAK47446.1                        | -                 |
| DR | PIR          | C70859    | C70859                            | -                 |
| DR | TIGR         | MT316     |                                   | -                 |
| DR | TubercuList  | Rv3032    |                                   | -                 |
| GO | GO           | 0016740   | F:transferase activity            | IEA               |
| DR | GO           | 0009058   | P:biosynthesis                    | IEA               |
| DR | InterPro     | IPR001296 | Glyco_trans_1                     |                   |
| DR | Pfam         | PF00534   | Glycos_transf_1                   |                   |
| DR | Hypothetical | protein   | Transferase                       | Complete proteome |
| DR | SO           | SEQUENCE  | 414 AA: 44805 MW: 64ADC74AE17A0DE | CR664             |



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.11727 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_362\_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | ID | Description        |
|------------|-------------|-------|--------|----|--------------------|
| 1          | 38          | 100.0 | 13     | 4  | ABBS6230 Vascular  |
| 2          | 38          | 100.0 | 13     | 4  | ABBS6225 Vascular  |
| 3          | 38          | 100.0 | 13     | 4  | AAU25032 Schizophr |
| 4          | 38          | 100.0 | 13     | 4  | AAU15376 Schizophr |
| 5          | 38          | 100.0 | 103    | 4  | AAU33145 Novel hum |
| 6          | 38          | 100.0 | 112    | 4  | AAO2636 Human pol  |
| 7          | 38          | 100.0 | 188    | 3  | AAU33145 Novel hum |
| 8          | 38          | 100.0 | 228    | 3  | AAU33145 Novel hum |
| 9          | 38          | 100.0 | 243    | 4  | AAU33087 Novel hum |
| 10         | 38          | 100.0 | 245    | 4  | AAU33081 Novel hum |
| 11         | 38          | 100.0 | 293    | 4  | AAU33082 Novel hum |
| 12         | 38          | 100.0 | 373    | 1  | AAU33087 Novel hum |
| 13         | 38          | 100.0 | 388    | 1  | AAU33087 Novel hum |
| 14         | 38          | 100.0 | 389    | 1  | AAU33087 Novel hum |
| 15         | 38          | 100.0 | 390    | 1  | AAU33087 Novel hum |
| 16         | 38          | 100.0 | 407    | 1  | AAU33087 Novel hum |
| 17         | 38          | 100.0 | 463    | 2  | AAU33087 Novel hum |
| 18         | 38          | 100.0 | 507    | 4  | AAU33086 Novel hum |
| 19         | 38          | 100.0 | 507    | 4  | AAU33086 Novel hum |
| 20         | 38          | 100.0 | 550    | 4  | AAU33086 Novel hum |
| 21         | 38          | 100.0 | 584    | 6  | ABG72381 Mature hu |
| 22         | 38          | 100.0 | 585    | 1  | ABG72381 Mature hu |
| 23         | 38          | 100.0 | 585    | 1  | ABG72381 Mature hu |
| 24         | 38          | 100.0 | 585    | 1  | ABG72381 Mature hu |
| 25         | 38          | 100.0 | 585    | 2  | AAU5318 Human ser  |

|    |    |       |     |   |                   |
|----|----|-------|-----|---|-------------------|
| 26 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 27 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 28 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 29 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 30 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 31 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 32 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 33 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 34 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 35 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 36 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 37 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 38 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 39 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 40 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 41 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 42 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 43 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 44 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |
| 45 | 38 | 100.0 | 585 | 2 | AAU5318 Human ser |

ALIGNMENTS

RESULT 1  
ID ABB56230 standard; peptide; 13 AA.  
XX ABB56230;  
AC ABB56230;  
XX  
DT 15-FEB-2002 (first entry)  
XX  
DE Vascular dementia-associated protein isoform (VPI) 430.  
XX  
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;  
XX diagnosis; prognosis; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200169261-A2.  
XX  
PD 20-SEP-2001.  
XX  
PF 14-MAR-2001; 2001WO-GB001106.  
XX  
PR 15-MAR-2000; 2000GB-00006285.  
PR 24-NOV-2000; 2000GB-00028734.  
PR 28-NOV-2000; 2000US-00724391.  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX Herath HMCAC, Parekh RB, Rohlf C;  
XX WPI; 2001-557937/62.  
XX  
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for  
PT determining stage of VD and monitoring the effect of VD therapy, for  
PT comprises analyzing body fluid by 2-dimensional electrophoresis for  
PT features correlated with VD.  
XX  
PS Claim 6; Page 39; 151pp; English.  
XX  
CC The invention relates to screening, diagnosis or prognosis of Vascular  
CC Dementia (VD) in a subject comprising analysis of body fluid from the  
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of  
CC features containing at least one chosen feature whose relative abundance  
CC correlates with the presence, absence, stage or severity of VD or  
CC predicts the onset or course of VD, especially detecting in a sample of  
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated  
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the  
CC specification. Detecting VD-associated features and VPI is useful for the  
CC screening, diagnosis or prognosis of VD, for determining the stage or

CC severity of VD, for identifying a subject at risk of VD or for monitoring  
 CC the effect of therapy administered to a subject having VD. Nucleic acids  
 CC encoding a VPI or inhibiting the function of a VPI are useful for the  
 CC treatment of VD and for gene therapy  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
 |||||  
 Db 3 AAADPHE 9

## RESULT 2

ABBS6225  
 ID ABB56225 standard; peptide; 13 AA.

XX  
 AC ABB56225;  
 XX  
 DT 15-FEB-2002 (first entry)  
 XX  
 DE Vascular dementia-associated protein isoform (VPI) 425.

XX Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;  
 KW diagnosis; prognosis; gene therapy.  
 KW  
 XX Homo sapiens.

XX WO200169261-A2.

XX 20-SEP-2001.

XX 14-MAR-2001; 2001WO-GB001106.

XX 15-MAR-2000; 2000GB-00006285.

XX 24-NOV-2000; 2000GB-00028734.

XX 28-NOV-2000; 2000US-00724391.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Herath HMAC, Parekh RB, Rohlf C;

XX WPI; 2001-557937/62.

XX Screening, diagnosis or prognosis of vascular dementia (VD), useful for  
 PT determining stage of VD and monitoring the effect of VD therapy,  
 PT comprises analyzing body fluid by 2-dimensional electrophoresis for  
 PT features correlated with VD.

XX Claim 6; Page 39; 151pp; English.

XX The invention relates to screening, diagnosis or prognosis of Vascular  
 CC Dementia (VD) in a subject comprising analysing body fluid from the  
 CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of  
 CC features containing at least one chosen feature whose relative abundance  
 CC correlates with the presence, absence, stage or severity of VD or  
 CC predicts the onset or course of VD, especially detecting in a sample of  
 CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated  
 CC protein isoforms (VPIs) (ABB5801-ABB56295) as fully defined in the  
 CC specification. Detecting VD-associated features and VPI is useful for the  
 CC screening, diagnosis or prognosis of VD, for determining the stage or  
 CC severity of VD, for identifying a subject at risk of VD or for monitoring  
 CC the effect of therapy administered to a subject having VD. Nucleic acids  
 CC encoding a VPI or inhibiting the function of a VPI are useful for the  
 CC treatment of VD and for gene therapy

XX Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAADPHE 7  
 |||||  
 Db 3 AAADPHE 9

## RESULT 3

AAU25032  
 ID AAU25032 standard; peptide; 13 AA.

XX AC AAU25032;

XX 18-DEC-2001 (first entry)

XX Schizophrenia-Associated Protein Isoform (SPI) peptide #261.

XX Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;  
 KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.

XX Homo sapiens.

XX WO200162785-A2.

XX 30-AUG-2001.

XX 23-FEB-2001; 2001WO-GB000792.

XX 24-FEB-2000; 2000GB-00004415.

XX 28-DEC-2000; 2000US-00750395.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Herath HMAC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;

XX WPI; 2001-570624/64.

XX New schizophrenia associated protein isoforms and encoding nucleic acid  
 PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia  
 PT and screening for potential drugs for treatment and new drug targets.

XX Disclosure; Page 34; 148pp; English.

XX The sequence represents a schizophrenia-associated protein isoform (SPI).  
 CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable  
 CC in cerebrospinal fluid, serum or plasma and are useful markers of  
 CC schizophrenia. The sequences can be used for treatment and diagnosis of  
 CC schizophrenia. Screening, prognosis, monitoring the results of therapy,  
 CC identifying patients most likely to respond to a particular therapy and  
 CC identification of new targets for drug treatment. SPI DNA is useful as a  
 CC nucleic acid probe to detect the presence of nucleic acids or SPIs

XX Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
 |||||  
 Db 3 AAADPHE 9

## RESULT 4

AAU15376  
 ID AAU15376 standard; peptide; 13 AA.

XX AC AAU15376;

XX 24-OCT-2001 (first entry)

XX Schizophrenia-associated isoform peptide #261.

KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;  
 KW neurological disorder; neuropathy.  
 XX Homo sapiens.  
 XX WO200163293-A2.  
 XX 30-AUG-2001.  
 XX 23-FEB-2001; 2001WO-GB000783.  
 XX 24-FEB-2000; 2000GB-00004415.  
 XX 28-DEC-2000; 2000US-00750395.  
 XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.  
 XX Herath HMAC, Parekh RB, Rohlf C;  
 XX WPI; 2001-502868/55.  
 XX Diagnosing and monitoring Schizophrenia by detecting the presence of  
 PT Schizophrenia Associated Features and Schizophrenia Associated Protein  
 PT isoforms in samples of cerebrospinal fluid.  
 XX Claim 6; Page 34; 160pp; English.  
 XX The invention relates to methods and compositions for screening,  
 CC diagnosis and prognosis of Schizophrenia. The method involves detecting  
 CC the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH  
 CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,  
 CC immunoblotting or hybridisation assay, for diagnosing and monitoring SCH,  
 CC studying the effectiveness of treatments and for identifying potential  
 CC therapeutic agents. The method is used for (1) screening or diagnosis of  
 CC SCH and the relative abundance of at least 1 chosen feature correlates  
 CC with the presence or absence of SCH; and (2) monitoring the effect of  
 CC therapy administered to a subject with SCH and the relative abundance of  
 CC at least 1 chosen feature which correlates with the severity of SCH. The  
 CC expression and activity of the SFs, SPIs and related molecules (e.g.  
 CC secondary messengers) are studied to diagnose SCH, monitor the progress  
 CC of the disorder and the effectiveness of treatment and as targets to  
 CC identify and produce potential therapeutic agents for the treatment of  
 CC SCH. The paucity of detectable neurologic defects distinguishes  
 CC neuropsychiatric disorders such as SCH from neurological disorders, where  
 CC manifestations of anatomical and biochemical changes have been identified  
 CC in many cases. Consequently the identification and characterisation of  
 CC cellular and/or molecular causative defects and neuropathies are  
 CC necessary for improved treatment of neuropsychiatric disorders. AAU5114-  
 CC AAU5762 represent the amino acid sequences of schizophrenia-associated  
 CC isoforms used in the method of the invention  
 XX SQ Sequence 13 AA;  
 Query Match 100.0%; Score 38; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAADPHE 7  
 |||||  
 Db 3 AAADPHE 9  
 RESULT 5  
 AAU33145  
 ID AAU33145 standard; protein; 103 AA.  
 XX AAU33145;  
 XX 18-DEC-2001 (first entry)  
 XX Novel human secreted protein #3636.  
 XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX Homo sapiens.  
 XX WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008656.  
 XX 18-APR-2000; 2000US-00552929.  
 XX 26-JAN-2001; 2001US-00770160.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX Claim 20; Page 712; 765pp; English.  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX SQ Sequence 103 AA;  
 Query Match 100.0%; Score 38; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAADPHE 7  
 |||||  
 Db 34 AAADPHE 40  
 RESULT 6  
 AA02636  
 ID AA02636 standard; protein; 112 AA.  
 XX AA02636;  
 XX 06-NOV-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 16528.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 XX Homo sapiens.  
 XX WO200164835-A2.  
 XX 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.  
 XX PR 28-FEB-2000; 2000US-00515126.  
 XX PR 18-MAY-2000; 2000US-00577409.  
 XX (HYSE-) HYSEQ INC.  
 XX PA Tang YT, Liu C, Drmanac RT;  
 XX PI WPI; 2001-514838/56.  
 XX DR N-PSDB; AAI82567.  
 XX DR Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX PS Claim 20; SEQ ID NO 16528; 1399pp + Sequence Listing; English.  
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 112 AA;  
 SQ

Query Match 100.0%; Score 38; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 5.9; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

OY 1 AAADPHE 7  
 DB 85 AAADPHE 91  
 |||||

RESULT 7  
 AAY83948  
 ID AAY83948 standard; protein; 188 AA.  
 AC AAY83948;  
 XX 28-JUL-2000 (first entry)  
 DT Yeast codon-biased recombinant HSA protein fragment HSA-II.  
 DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
 KW overlapping oligonucleotide; expression vector.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX CN1239103-A.  
 PN 22-DEC-1999.  
 XX 17-JUN-1998; 98CN-00102506.  
 PF 17-JUN-1998; 98CN-00102506.  
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
 PA Li S, Lu D;  
 PI WPI; 2000-351198/31.  
 XX N-PSDB; AAI10093.

XX PT Process for preparing recombinant human serum albumin comprising yeast  
 PT biased sex codons - uses a recombinant DNA technique.  
 XX Example 1; Fig 5; 44pp; Chinese.  
 XX The method relates to a method of recombinantly producing human serum  
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
 CC a yeast codon bias. The complete HSA gene (AAAI0091) was generated as  
 CC three synthetic fragments (AAAI0092-AI0094) joined by recombinant DNA  
 CC technology. Each HSA fragment was synthesised from overlapping  
 CC oligonucleotide fragments that were extended. This sequence represents  
 CC the sequence of the HSA fragment HSA-II encoded by the human gene with a  
 CC yeast codon bias. The invention also covers a recombinant expression and  
 CC vector, yeast host cells carrying the recombinant expression vector and  
 CC the process for producing human serum albumin in the yeast host cell,  
 CC especially in secretory mode  
 XX Sequence 188 AA;  
 SQ

Query Match 100.0%; Score 38; DB 3; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 10; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

OY 1 AAADPHE 7  
 DB 178 AAADPHE 184  
 |||||

RESULT 8  
 AAY83949  
 ID AAY83949 standard; protein; 228 AA.  
 XX AAY83949;  
 AC AAY83949;  
 XX 28-JUL-2000 (first entry)  
 DT Yeast codon-biased recombinant HSA protein fragment HSA-III.  
 DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
 KW overlapping oligonucleotide; expression vector.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX CN1239103-A.  
 PN 22-DEC-1999.  
 XX 17-JUN-1998; 98CN-00102506.  
 PF 17-JUN-1998; 98CN-00102506.  
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
 PA Li S, Lu D;  
 PI WPI; 2000-351198/31.  
 XX N-PSDB; AAI10094.  
 XX Process for preparing recombinant human serum albumin comprising yeast  
 PT biased sex codons - uses a recombinant DNA technique.  
 XX Example 1; Fig 7; 44pp; Chinese.  
 XX The method relates to a method of recombinantly producing human serum  
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
 CC a yeast codon bias. The complete HSA gene (AAAI0091) was generated as  
 CC three synthetic fragments (AAAI0092-AI0094) joined by recombinant DNA  
 CC technology. Each HSA fragment was synthesised from overlapping  
 CC oligonucleotide fragments that were extended. This sequence represents  
 CC the sequence of the HSA fragment HSA-III encoded by the human gene with a  
 CC yeast codon bias. The invention also covers a recombinant expression

CC vector, yeast host cells carrying the recombinant expression vector and  
CC the process for producing human serum albumin in the yeast host cell,  
CC especially in secretory mode  
XX  
SQ Sequence 228 AA;

Query Match 100.0%; Score 38; DB 3; Length 228;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7  
Db 5 AAADPHE 11  
|||||

## RESULT 9

AAU33087  
ID AAU33087 standard; protein; 243 AA.

XX AC AAU33087;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #3578.

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

XX 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
FT vaccination, testing and therapy.

XX Claim 20; Page 706; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention

XX Sequence 243 AA;

Query Match 100.0%; Score 38; DB 4; Length 243;  
Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAADPHE 7  
Db 64 AAADPHE 70  
|||||

## RESULT 10

AAU33081  
ID AAU33081 standard; protein; 245 AA.

XX AC AAU33081;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #3572.

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

XX 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
FT vaccination, testing and therapy.

XX Claim 20; Page 705; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention

XX Sequence 245 AA;

Query Match 100.0%; Score 38; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7  
Db 15 AAADPHE 21  
|||||

## RESULT 11

|    |  |  |
|----|--|--|
| XX | N-terminal of human serum albumin polypeptide.                           |  |
| DE |  |  |
| XX | Human serum albumin polypeptide; plasma expanders.                       |  |
| XX |  |  |
| OS | Homo sapiens; (Human).   |  |
| XX |  |  |
| PN | EP322094-A.  |  |
| XX |  |  |
| PD | 28-JUN-1989.   |  |
| XX |  |  |
| PF | 25-OCT-1988; 88EP-00310000.  |  |
| XX |  |  |
| PR | 30-OCT-1987; 87GB-00025529.  |  |
| XX |  |  |
| PA | (DELZ ) DELTA BIOTECHNOLOGY LTD.   |  |
| XX |  |  |
| PI | Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;                         |  |
| XX |  |  |
| DR | WPI; 1989-186464/26.   |  |
| XX |  |  |
| XX | New N-terminal fragments of human serum albumin - esp. useful as blood   |  |
| PT | plasma expanders.  |  |
| PS | Claim 2; Page 9; 20pp; English.  |  |
| XX |  |  |
| CC | N-terminal portion of human serum albumin. Used as plasma expanders, or  |  |
| CC | as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-  |  |
| CC | MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS |  |
| CC | field)   |  |
| XX |  |  |
| SQ | Sequence 373 AA;   |  |
|    |  |  |
|    | Query Match 100.0%; Score 38; DB 1; Length 373;                          |  |
|    | Best Local Similarity 100.0%; Pred. NO. 21;                              |  |
|    | Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0                |  |
| QY | 1 AAADPHE 7  |  |
|    |  |  |
| DB | 362 AAADPHE 368  |  |
|    |  |  |
|    | RESULT 13  |  |
|    | AAP90389   |  |
| ID | AAP90389 standard; protein; 388 AA.                                      |  |
| XX |  |  |
| AC | AAP90389;  |  |
| XX |  |  |
| DT | 24-OCT-2003 (revised)  |  |
| DT | 25-MAR-2003 (revised)  |  |
| DT | 01-NOV-1989 (first entry)  |  |
| XX |  |  |
| DE | N-terminal human serum albumin polypeptide.                              |  |
| XX |  |  |
| KW | N-terminal human serum albumin polypeptide; plasma expanders.            |  |
| XX |  |  |
| OS | Homo sapiens; (Human).   |  |
| XX |  |  |
| PN | EP322094-A.  |  |
| XX |  |  |
| PD | 28-JUN-1989.   |  |
| XX |  |  |
| PF | 25-OCT-1988; 88EP-00310000.  |  |
| XX |  |  |
| PR | 30-OCT-1987; 87GB-00025529.  |  |
| XX |  |  |
| PA | (DELZ ) DELTA BIOTECHNOLOGY LTD.   |  |
| XX |  |  |
| PI | Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;                         |  |
| XX |  |  |
| DR | WPI; 1989-186464/26.   |  |
| XX |  |  |
| PT | New N-terminal fragments of human serum albumin - esp. useful as blood   |  |

```

PT plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
XX
CC N-terminal fragment of human serum albumin used as plasma expander, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX Sequence 388 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 388;
    Best Local Similarity 100.0%; Pred. No. 22;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 362 AAADPHE 368
|||||

RESULT 14
AAP90390
ID AAP90390 standard; protein; 389 AA.
XX
XX AAP90390;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 01-NOV-1989 (first entry)
XX
XX N-terminal human serum albumin.
XX
XX N-terminal portion of human serum albumin; plasma expanders.
XX
XX Homo sapiens; (Human).
XX
XX EP322094-A.
XX
XX 28-JUN-1989.
XX
XX 25-OCT-1988; 88EP-00310000.
XX
XX 30-OCT-1987; 87GB-00025529.
XX
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
XX WPI; 1989-186464/26.
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
XX
XX N-terminal portion of human serum albumin. Used to make plasma expanders,
XX or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
XX -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 389 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 389;
    Best Local Similarity 100.0%; Pred. No. 22;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 362 AAADPHE 368
|||||

RESULT 15
AAP90391

```

```

ID AAP90391 standard; protein; 390 AA.
XX
XX AAP90391;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 01-NOV-1989 (first entry)
XX
XX N-terminal human serum albumin.
XX
XX N-terminal portion of human serum albumin; plasma expanders.
XX
XX Homo sapiens; (Human).
XX
XX EP322094-A.
XX
XX 28-JUN-1989.
XX
XX 25-OCT-1988; 88EP-00310000.
XX
XX 30-OCT-1987; 87GB-00025529.
XX
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
XX WPI; 1989-186464/26.
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
XX
XX N-terminal portion of human serum albumin. Used to make new N-terminal
XX fragments which are used as plasma expanders, or as substitutes for HSA
XX or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA
XX field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 390 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 390;
    Best Local Similarity 100.0%; Pred. No. 23;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 362 AAADPHE 368
|||||

Search completed: April 19, 2004, 11:51:20
Job time : 4.11727 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.94737 Seconds

(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_362\_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 38    | 100.0       | 13     | 9  | US-09-791-378-261 |
| 2          | 38    | 100.0       | 585    | 9  | US-09-929-552-2   |
| 3          | 38    | 100.0       | 585    | 10 | US-09-932-613-445 |
| 4          | 38    | 100.0       | 585    | 10 | US-09-984-010-26  |
| 5          | 38    | 100.0       | 585    | 10 | US-09-833-041-18  |
| 6          | 38    | 100.0       | 585    | 10 | US-09-833-117-18  |
| 7          | 38    | 100.0       | 585    | 10 | US-09-932-322-445 |
| 8          | 38    | 100.0       | 585    | 10 | US-09-832-501-18  |
| 9          | 38    | 100.0       | 585    | 11 | US-09-833-118-18  |
| 10         | 38    | 100.0       | 585    | 11 | US-09-833-245-18  |
| 11         | 38    | 100.0       | 585    | 12 | US-10-424-999-11  |
| 12         | 38    | 100.0       | 585    | 12 | US-10-425-000-31  |
| 13         | 38    | 100.0       | 585    | 12 | US-10-433-108-34  |
| 14         | 38    | 100.0       | 585    | 13 | US-10-153-064-5   |
| 15         | 38    | 100.0       | 585    | 14 | US-10-153-064A-5  |

|    |    |       |     |    |                   |                    |
|----|----|-------|-----|----|-------------------|--------------------|
| 16 | 38 | 100.0 | 585 | 14 | US-10-319-263-1   | Sequence 1, Appli  |
| 17 | 38 | 100.0 | 585 | 14 | US-10-319-263-2   | Sequence 2, Appli  |
| 18 | 38 | 100.0 | 585 | 14 | US-10-414-469-1   | Sequence 1, Appli  |
| 19 | 38 | 100.0 | 585 | 14 | US-10-414-469-2   | Sequence 2, Appli  |
| 20 | 38 | 100.0 | 585 | 14 | US-10-413-831-1   | Sequence 1, Appli  |
| 21 | 38 | 100.0 | 585 | 14 | US-10-413-831-2   | Sequence 2, Appli  |
| 22 | 38 | 100.0 | 585 | 15 | US-10-413-832-1   | Sequence 1, Appli  |
| 23 | 38 | 100.0 | 585 | 15 | US-10-413-832-2   | Sequence 2, Appli  |
| 24 | 38 | 100.0 | 585 | 15 | US-10-414-386-1   | Sequence 1, Appli  |
| 25 | 38 | 100.0 | 585 | 15 | US-10-414-386-2   | Sequence 2, Appli  |
| 26 | 38 | 100.0 | 585 | 15 | US-10-233-675A-11 | Sequence 11, Appli |
| 27 | 38 | 100.0 | 585 | 15 | US-10-462-262-26  | Sequence 26, Appli |
| 28 | 38 | 100.0 | 604 | 10 | US-09-984-010-7   | Sequence 7, Appli  |
| 29 | 38 | 100.0 | 609 | 10 | US-09-919-039-370 | Sequence 370, App  |
| 30 | 38 | 100.0 | 609 | 12 | US-10-609-346-12  | Sequence 12, Appli |
| 31 | 38 | 100.0 | 609 | 13 | US-10-153-064-7   | Sequence 7, Appli  |
| 32 | 38 | 100.0 | 609 | 14 | US-10-153-604A-7  | Sequence 23, Appli |
| 33 | 38 | 100.0 | 609 | 14 | US-10-365-623-23  | Sequence 2, Appli  |
| 34 | 38 | 100.0 | 610 | 9  | US-09-984-186-2   | Sequence 2, Appli  |
| 35 | 38 | 100.0 | 610 | 14 | US-10-237-667-2   | Sequence 2, Appli  |
| 36 | 38 | 100.0 | 610 | 14 | US-10-237-708-2   | Sequence 2, Appli  |
| 37 | 38 | 100.0 | 610 | 14 | US-10-237-866-2   | Sequence 2, Appli  |
| 38 | 38 | 100.0 | 610 | 14 | US-10-237-871-2   | Sequence 2, Appli  |
| 39 | 38 | 100.0 | 610 | 14 | US-10-237-824-2   | Sequence 13, Appli |
| 40 | 38 | 100.0 | 616 | 12 | US-10-433-108-13  | Sequence 16, Appli |
| 41 | 38 | 100.0 | 624 | 12 | US-10-433-108-16  | Sequence 14, Appli |
| 42 | 38 | 100.0 | 631 | 12 | US-10-433-108-14  | Sequence 15, Appli |
| 43 | 38 | 100.0 | 640 | 12 | US-10-433-108-15  | Sequence 17, Appli |
| 44 | 38 | 100.0 | 640 | 12 | US-10-433-108-17  | Sequence 133, App  |
| 45 | 38 | 100.0 | 651 | 13 | US-10-153-064-133 |                    |

#### ALIGNMENTS

#### RESULT 1

US-09-791-378-261  
; Sequence 261, Application US/09791378  
; Patent No. US20020142303A1  
; GENERAL INFORMATION:  
; APPLICANT: Parekh, Rajesh  
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF SCHIZOPHRENIA  
; FILE REFERENCE: 9195-061-399  
; CURRENT APPLICATION NUMBER: US/09/791,378  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 09/750,395  
; PRIOR FILING DATE: 2000-12-28  
; NUMBER OF SEQ ID NOS: 677  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 261  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-378-261

Query Match 100.0%; Score 38; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 1 AAADPHE 7

Db 3 AAADPHE 9

#### RESULT 2

US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Somnenschein, Carlos  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells



NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/929,552  
APPLICATION NUMBER: US/09/929,552  
FILING DATE: 14-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/769,746  
FILING DATE: 19-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: MBRI-02584  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-929-552-2

Query Match 100.0%; Score 38; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
|||||  
Db 362 AAADPHE 368

RESULT 3  
US-09-932-613-445  
; Sequence 445, Application US/09932613  
; Publication No. US20030091565A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; APPLICANT: Beltzer, James P.  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Fleming, Tony J.  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
; FILE REFERENCE: DXX-025.1 PCT: DXX-025.1 US  
; CURRENT APPLICATION NUMBER: US/09/932,613  
; CURRENT FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 445  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-09-932-613-445

Query Match 100.0%; Score 38; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
|||||

Db 362 AAADPHE 368

RESULT 4  
US-09-984-010-26  
; Sequence 26, Application US/09984010  
; Publication No. US20030104578A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David James  
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE  
AND SERUM ALBUMIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION NUMBER: US/09/984,010  
; FILING DATE: 21-May-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/091,873  
; FILING DATE: 25-JUN-1998  
; APPLICATION NUMBER: PCT/GB96/03164  
; FILING DATE: 19-DEC-1996  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-984-010-26

Query Match 100.0%; Score 38; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
|||||  
Db 362 AAADPHE 368

RESULT 5  
US-09-833-041-18  
; Sequence 18, Application US/09833041  
; Publication No. US20030125247A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseitine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF545  
; CURRENT APPLICATION NUMBER: US/09/833,041  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 18

; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-041-18

Query Match 100.0%; Score 38; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
DB 362 AAADPHE 368

## RESULT 6

US-09-833-117-18  
; Sequence 18, Application US/09833117  
; Publication No. US20030171267A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF543  
; CURRENT APPLICATION NUMBER: US/09/833,117  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-117-18

Query Match 100.0%; Score 38; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
DB 362 AAADPHE 368

## RESULT 7

US-09-932-322-445  
; Sequence 445, Application US/09932322  
; Publication No. US20030194743A1  
; GENERAL INFORMATION:  
; APPLICANT: Dyax Corp.  
; APPLICANT: Beltzer, James P.  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Fleming, Tony J.  
; APPLICANT: Ladner, Robert Charles  
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (Blys)  
; FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US  
; CURRENT APPLICATION NUMBER: US/09/932,322  
; CURRENT FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 445  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-932-322-445

Query Match 100.0%; Score 38; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
DB 362 AAADPHE 368

## RESULT 8

US-09-832-501-18  
; Sequence 18, Application US/09832501  
; Publication No. US20030199043A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J.  
; APPLICANT: Sleep, Darrell  
; APPLICANT: Turner, Andrew J.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF542  
; CURRENT APPLICATION NUMBER: US/09/832,501  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-832-501-18

Query Match 100.0%; Score 38; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
DB 362 AAADPHE 368

## RESULT 9

US-09-833-118-18  
; Sequence 18, Application US/09833118  
; Publication No. US20030219875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF544  
; CURRENT APPLICATION NUMBER: US/09/833,118  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-118-18

Query Match 100.0%; Score 38; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 AAADPHE 7
Db      362 AAADPHE 368

RESULT 10
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
Db      362 AAADPHE 368

RESULT 11
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
Db      362 AAADPHE 368

RESULT 12
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
Db      362 AAADPHE 368

RESULT 13
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
Db      362 AAADPHE 368

RESULT 14
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
```

; PRIOR APPLICATION NUMBER: 60/293,212  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-10-153-064-5

Query Match 100.0%; Score 38; DB 13; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
 Db 362 AAADPHE 368

RESULT 15  
 US-10-153-604A-5  
 ; Sequence 5, Application US/10153604A  
 ; Publication No. US20030143191A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell et al.  
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
 ; FILE REFERENCE: PFS56  
 ; CURRENT APPLICATION NUMBER: US/10/153,604A  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/293,212  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-10-153-604A-5

Query Match 100.0%; Score 38; DB 14; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7  
 Db 362 AAADPHE 368

Search completed: April 19, 2004, 12:54:59  
 Job time : 2.94737 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_280\_288

Perfect score: 47  
Sequence: 1 EKPLLEKSH 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Match | Length | ID     | Description         |
|------------|-------|-------|--------|--------|---------------------|
| 1          | 47    | 100.0 | 417    | Q86YGO | Q86YGO homo sapien  |
| 2          | 39    | 83.0  | 584    | Q7YSG3 | Q7YSG3 felis silve  |
| 3          | 37    | 78.7  | 245    | Q7WQW3 | Q7WQW3 bordetella   |
| 4          | 37    | 78.7  | 245    | Q7WYI5 | Q7WYI5 bordetella   |
| 5          | 37    | 78.7  | 576    | Q8C7C7 | Q8C7C7 mus musculu  |
| 6          | 37    | 78.7  | 608    | Q95VB7 | Q95VB7 schistosoma  |
| 7          | 37    | 78.7  | 608    | Q8C7H3 | Q8C7H3 mus musculu  |
| 8          | 37    | 78.7  | 1381   | Q94C01 | Q94C01 hordeum vul  |
| 9          | 36    | 76.6  | 131    | Q8LFB3 | Q8LFB3 arabidopsis  |
| 10         | 36    | 76.6  | 480    | Q7UQ53 | Q7UQ53 rhodospirill |
| 11         | 35    | 74.5  | 123    | Q8KLA6 | Q8KLA6 rhizobium e  |
| 12         | 35    | 74.5  | 156    | Q97F20 | Q97F20 clostridium  |
| 13         | 35    | 74.5  | 226    | Q8I621 | Q8I621 plasmodium   |
| 14         | 35    | 74.5  | 301    | Q9PMW9 | Q9PMW9 campylobact  |
| 15         | 35    | 74.5  | 1265   | Q01495 | Q01495 caenorhabdi  |
| 16         | 35    | 74.5  | 1380   | Q40001 | Q40001 hordeum vul  |

|    |    |      |      |    |        |                     |
|----|----|------|------|----|--------|---------------------|
| 17 | 35 | 74.5 | 1534 | 5  | Q8MPV7 | Q8MPV7 caenorhabdi  |
| 18 | 35 | 74.5 | 1538 | 5  | Q8MPV6 | Q8MPV6 caenorhabdi  |
| 19 | 34 | 72.3 | 138  | 2  | Q84B18 | Q84B18 streptococc  |
| 20 | 34 | 72.3 | 191  | 16 | Q81SB7 | Q81SB7 bacillus an  |
| 21 | 34 | 72.3 | 191  | 16 | Q81FA3 | Q81FA3 bacillus ce  |
| 22 | 34 | 72.3 | 226  | 16 | Q8ZSI9 | Q8ZSI9 anabaena sp  |
| 23 | 34 | 72.3 | 244  | 2  | Q34Z50 | Q34Z50 wolinnella s |
| 24 | 34 | 72.3 | 250  | 4  | Q9NS78 | Q9NS78 homo sapien  |
| 25 | 34 | 72.3 | 265  | 4  | Q9NS79 | Q9NS79 homo sapien  |
| 26 | 34 | 72.3 | 269  | 4  | Q3H4F5 | Q3H4F5 homo sapien  |
| 27 | 34 | 72.3 | 269  | 11 | Q9CQ04 | Q9CQ04 mus musculu  |
| 28 | 34 | 72.3 | 269  | 11 | Q9DAD7 | Q9DAD7 mus musculu  |
| 29 | 34 | 72.3 | 311  | 2  | Q68651 | Q68651 haemophilus  |
| 30 | 34 | 72.3 | 477  | 16 | Q87Y51 | Q87Y51 pseudomonas  |
| 31 | 34 | 72.3 | 528  | 10 | Q94FY3 | Q94FY3 zea mays (m  |
| 32 | 34 | 72.3 | 560  | 13 | Q7ZTT2 | Q7ZTT2 brachydanio  |
| 33 | 34 | 72.3 | 586  | 3  | Q9P985 | Q9P985 saccharomyc  |
| 34 | 34 | 72.3 | 586  | 3  | Q9URE1 | Q9URE1 saccharomyc  |
| 35 | 34 | 72.3 | 586  | 3  | Q9P986 | Q9P986 saccharomyc  |
| 36 | 34 | 72.3 | 586  | 3  | Q9P984 | Q9P984 saccharomyc  |
| 37 | 34 | 72.3 | 586  | 3  | Q8J2S8 | Q8J2S8 saccharomyc  |
| 38 | 34 | 72.3 | 747  | 5  | Q7YSP6 | Q7YSP6 caenorhabdi  |
| 39 | 34 | 72.3 | 768  | 2  | Q9RLB8 | Q9RLB8 ruminococcu  |
| 40 | 34 | 72.3 | 861  | 5  | Q3XWF9 | Q3XWF9 caenorhabdi  |
| 41 | 33 | 70.2 | 171  | 16 | Q8ZQMO | Q8ZQMO salmonella   |
| 42 | 33 | 70.2 | 173  | 16 | Q98KG5 | Q98KG5 rhizobium 1  |
| 43 | 33 | 70.2 | 195  | 4  | Q9H855 | Q9H855 homo sapien  |
| 44 | 33 | 70.2 | 270  | 16 | Q7VK53 | Q7VK53 helicobacte  |
| 45 | 33 | 70.2 | 271  | 16 | Q8XLY6 | Q8XLY6 clostridium  |

## ALIGNMENTS

## RESULT 1

Q86YGO PRELIMINARY; PRT; 417 AA.  
AC Q86YGO;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to alpha-fetoprotein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (JSC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041789; AAH41789.1; -  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot.; 2.  
DR PRINTS; PR00302; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 47; DB 4; Length 417;

Best Local Similarity 100.0%; Pred. No. 0.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9

Db 112 EKPLLEKSH 120

## RESULT 2

```

Q7YSG3
ID Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD22275.1; -.
FT NON_TER
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
Query Match 83.0%; Score 39; DB 16; Length 584;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KPLLEKSH 9
Db 281 KPVLEKSH 288

RESULT 3
Q7WQW3 PRELIMINARY; PRT; 245 AA.
ID Q7WQW3
AC Q7WQW3
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Phosphorylase family protein.
GN BB0212.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30710.1; -.
KW Complete proteome.
SQ SEQUENCE 245 AA; 25969 MW; 6655450B3B72FB97 CRC64;

Query Match 78.7%; Score 37; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 110 EYPLLEKSH 118

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7
AC Q8C7C7
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573(2002).
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.

```

```

Db 110 EYPLLEKSH 118
RESULT 4
Q7WLY5 PRELIMINARY; PRT; 245 AA.
ID Q7WLY5
AC Q7WLY5
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Phosphorylase family protein.
GN BPP0208.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12922 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE39949.1; -.
KW Complete proteome.
SQ SEQUENCE 245 AA; 25999 MW; 3B2924005602E7EE CRC64;

Query Match 78.7%; Score 37; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 110 EYPLLEKSH 118

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7
AC Q8C7C7
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573(2002).
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.

```

```
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 576 AA; 65002 MW; F85733B99AE37F04 CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 11; Length 576;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:||||:|
Db 272 DKPLLEKSH 280

RESULT 6
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., Loverde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF18550; AL08579.1; -.
DR GO: GO:0006810; P:transport; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E54 CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 5; Length 608;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:||||:|
Db 304 DKPLLEKSH 312

RESULT 7
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP
```

```
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:187991; Alb1.
DR GO: GO:0005815; C:extracellular space; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 11; Length 608;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:||||:|
Db 304 DKPLLEKSH 312

RESULT 8
Q94C01 PRELIMINARY; PRT; 1381 AA.
ID Q94C01;
AC Q94C01;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mg-chelataase subunit XANTHA-F.
GN XANTHA-F.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Svalcof's Bonus;
RA Olsson U.;
RT "Barley (Hordeum vulgare) Mg-chelataase subunit (Xantha-f) gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY039003; AAK72401.1; -.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR InterPro: IPR003672; CobN/Mg_chltase.
DR Pfam; PF02514; cobN-Mg_chel; 1.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 1381 AA; 152900 MW; ED91EA6CFOF23B5B CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 10; Length 1381;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KPLLEKSH 9
:||||:|
Db 465 KPLLEKSH 472

RESULT 9
Q8LFB3 PRELIMINARY; PRT; 191 AA.
ID Q8LFB3;
AC Q8LFB3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation."; 0:0-0(2002).  
 RL Genome Biol. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RT "Full-length cDNA from Arabidopsis thaliana";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY084948; AAM61509.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 191 AA; 21908 MW; F7434F97C7520294 CRC64;  
 Query Match 76.6%; Score 36; DB 10; Length 191;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 EKPLEKSH 9  
 Db 151 EKPLEKKN 159  
 RESULT 10  
 Q7UQ53 PRELIMINARY; PRT; 480 AA.  
 AC Q7UQ53;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN RB6510.  
 OS Rhodospirillum rubrum.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Firellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Bortym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Firellula sp.  
 RT strain 1."; 0:0-0(2003).  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL: BX294144; CAD74852.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 480 AA; 53002 MW; 82507EAS798FB7A2 CRC64;  
 Query Match 76.6%; Score 36; DB 16; Length 480;  
 Best Local Similarity 87.5%; Pred. No. 91;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 KPLEKSH 9  
 Db 66 KPLEKHL 73  
 RESULT 11  
 Q8KLA6 PRELIMINARY; PRT; 123 AA.  
 AC Q8KLA6;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein.  
 GN YP017.  
 OS Rhizobium etli.  
 OC Plasmid symbiotic plasmid P42d  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=29449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RX MEDLINE=91193195; PubMed=2013564;  
 RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;  
 RT "Structural complexity of the symbiotic plasmid of Rhizobium  
 RT leguminosarum bv. phaseoli."; 0:0-0(1991).  
 RL J. Bacteriol. 173:2411-2419(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RX MEDLINE=97419521; PubMed=9274036;  
 RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,  
 RA Cevallos M.A., Davila G.;  
 RT "Sequence, localization and characteristics of the replicator region  
 RT of the symbiotic plasmid of Rhizobium etli."; 0:0-0(1997).  
 RL Microbiology 143:2825-2831(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RA Quintero V., Cevallos M.A., Davila G.;  
 RT "A site-specific recombinase and RecA are required to exert  
 RT incompatibility towards the symbiotic plasmid of Rhizobium etli."; 0:0-0(1997).  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U80928; AAM54841.1; -;  
 DR GO: GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; plasmid.  
 SQ SEQUENCE 123 AA; 13938 MW; 7349C06DC4D4766A CRC64;  
 Query Match 74.5%; Score 35; DB 2; Length 123;  
 Best Local Similarity 75.0%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 KPLEKSH 9  
 Db 38 KPLEKSH 45  
 RESULT 12  
 Q97F20 PRELIMINARY; PRT; 156 AA.  
 AC Q97F20;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Predicted acetyltransferase.  
 GN CAC2935.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466286;  
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium Clostridium acetobutylicum"; 0:0-0(2001).  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE007790; AAK80877.1; -;  
 DR FIR; B97261; B97261.  
 DR GO: GO:0008080; F:N-acetyltransferase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.



```

DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf; 1.
KW Transferrase; Complete proteome.
SQ SEQUENCE 156 AA; 17998 MW; ADF750DFC5953E8 CRC64;

Query Match
Best Local Similarity 74.5%; Score 35; DB 16; Length 156;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9
    |||||
Db 137 EKPLEIETH 145

RESULT 13
ID Q81621 PRELIMINARY; PRT; 226 AA.
AC Q81621;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cyclophilin, putative.
GN PF0120C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pitera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum".
RL Nature 419:498-511(2002).
DR EMBL; AE014844; AAN36113.1; -.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR002130; CSA_PPIase.
DR Pfam; PF00160; pro_isomerase; 1.
DR PRINTS; PR00153; CSAPPISMRASE.
DR PROSITE; PS50072; CSA_PPIase_2; 1.
SQ SEQUENCE 226 AA; 26429 MW; 3A5E790BB457E181 CRC64;

Query Match
Best Local Similarity 74.5%; Score 35; DB 5; Length 226;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEK 7
    |||||
Db 197 EKPLEEK 203

RESULT 14
ID Q9PMW9 PRELIMINARY; PRT; 301 AA.
AC Q9PMW9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Cj1330.
GN Cj1330.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139078; CAB73757.1; -.
DR PIR; G81276; G81276.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 301 AA; 35002 MW; 3A14BE93BA6F61D9 CRC64;

Query Match
Best Local Similarity 74.5%; Score 35; DB 16; Length 301;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EKPILLEKS 8
    |||||
Db 86 EKPLFEKS 93

RESULT 15
ID O01495 PRELIMINARY; PRT; 1265 AA.
AC O01495;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C34G6.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Langston Y., Rohlfing T.;
RT "The sequence of C. elegans cosmid C34G6."
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; U97407; AAB52482.2; -.
DR PIR; D87789; D87789.
DR WormPep; C34G6.4; CE29212.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD0000006; ABC_transporter; 2.

```

Mon Apr 19 13:27:19 2004

DR SMART; SM00382; AAA; 2.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
KW Hypothetical protein; ATP-binding; Transport  
SQ SEQUENCE 1265 AA; 140465 MW; 4948EF5C5A402757 CRC64;  
Query Match 74.5%; Score 35; DB 5; Length 1265;  
Best Local Similarity 75.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KPLEKSH 9  
Db 13 KPLKRS 20

Search completed: April 19, 2004, 12:00:09  
Job time : 5.47368 secs

DR InterPro; IPR006140; 2-Hacid DH\_C.  
 DR Pfam; PF00389; 2-Hacid DH; 1.  
 DR Pfam; PF02826; 2-Hacid DH\_C; 1.  
 KW plasmid; Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 315 AA; 34163 MW; A3357B4D7608BE26 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 315;  
 Best Local Similarity 85.7%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 :|||||  
 Db 31 EAADKAA 37

RESULT 11  
 Q7U021 PRELIMINARY; PRT; 387 AA.  
 ID Q7U021;  
 AC Q7U021;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Probable succinyl-CoA synthetase (Beta chain) SUCC (SCS-beta)  
 DE (EC 6.2.1.5).  
 DE (EC 6.2.1.5).  
 GN SUCC OR MB0976.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RA "The complete genome sequence of Mycobacterium bovis."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
 DR EMBL; BX248337; CAD93837.1;  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 387 AA; 40897 MW; DC055A7CD5E987CC CRC64;

Query Match 90.6%; Score 29; DB 16; Length 387;  
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 :|||||  
 Db 376 EAADKAA 382

RESULT 12  
 Q82PD6 PRELIMINARY; PRT; 466 AA.  
 ID Q82PD6;  
 AC Q82PD6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN SAV987.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=2447403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Korioka H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RP "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis; deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RA "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005025; BAC68697.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 466 AA; 48152 MW; 436EEF5AC2E3FD2 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 466;  
 Best Local Similarity 85.7%; Pred. No. 4.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 :|||||  
 Db 60 EAADKAA 66

RESULT 13  
 Q8NST2 PRELIMINARY; PRT; 503 AA.  
 ID Q8NST2;  
 AC Q8NST2;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein Cg10585.  
 GN CGL0585.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005275; BAB97978.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 503 AA; 52191 MW; 7C257A5457A515C1 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 503;  
 Best Local Similarity 85.7%; Pred. No. 5.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 :|||||  
 Db 106 EAADKAA 112

RESULT 14  
 Q8XN8 PRELIMINARY; PRT; 574 AA.  
 ID Q8XN8;  
 AC Q8XN8;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE LD23509P (CG12141-PA).  
 GN AATS-LYS OR CG12141.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.,  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides F.G., Scher S.E., Li P.W., Hoskins R.A., Calle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Aril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Hariri N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weissstock G.M., Weissbach J.,  
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jallali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RT "Sequencing of *Drosophila melanogaster* genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,

RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RT "Annotation of *Drosophila melanogaster* genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY089547; AL90285.1; -  
DR EMBL; AE003447; AA009255.1; -  
DR FlyBase; FB00027084; Aats-lys.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004824; F:lysine-tRNA ligase activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0006430; P:lysyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR008994; Nucleic acid OB.  
DR InterPro; IPR004364; tRNA-synt\_2.  
DR InterPro; IPR002313; tRNA-synt\_lys\_2.  
DR InterPro; IPR004365; tRNA-anti.  
DR InterPro; IPR006195; tRNA\_ligase\_II.  
DR Pfam; PF00152; tRNA-synt\_2; 1.  
DR Pfam; PF01336; tRNA-anti; 1.  
DR PRINTS; PR00982; TRNASYNTHLYS.  
DR TIGRFAMs; TIGR00499; lysS\_bact; 1.  
DR PROSITE; PS00862; AA TRNA\_LIGASE\_II; 1.  
SQ SEQUENCE 574 AA; 64661 MW; 861DAEA9C53E0DBE CRC64;  
Query Match 90.6%; Score 29; DB 5; Length 574;  
Best Local Similarity 85.7%; Pred. No. 5.9e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
DB 498 QASDKAA 504  
RESULT 15  
QYSG3 PRELIMINARY; PRT; 584 AA.  
ID QYSG3  
AC QYSG3  
DT 01-OCT-2003 (TReMBLrel. 25, Created)  
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Albumin (Fragment).  
GN ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=livar;  
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,  
RA Rumpold H., Valenta R., Spitzauer S.ification of recombinant cat  
RT "Escherichia coli expression and purification of recombinant cat  
RT albumin.IgE recognition, induction of basophil activation and  
RT lymphoproliferative responses in atopic patients."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ487677; CAD32275.1; -  
FT NON\_TER 1  
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;  
Query Match 90.6%; Score 29; DB 6; Length 584;  
Best Local Similarity 85.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
DB 498 QASDKAA 504

Db 170 EAADKAA 176

Search completed: April 19, 2004, 12:00:05  
Job time : 4.70175 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.11727 Seconds

(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_170\_176

Perfect score: 32

Sequence: 1 QADKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 32    | 100.0       | 116    | 4     | AAO00108 Human pol |
| 2          | 32    | 100.0       | 143    | 4     | AAO02571 Human pol |
| 3          | 32    | 100.0       | 192    | 4     | AAU29875 Novel hum |
| 4          | 32    | 100.0       | 195    | 5     | AAO17048 Human ser |
| 5          | 32    | 100.0       | 195    | 7     | ABU10022 Human ser |
| 6          | 32    | 100.0       | 204    | 3     | AAy83947 Yeast cod |
| 7          | 32    | 100.0       | 214    | 4     | AAU29874 Novel hum |
| 8          | 32    | 100.0       | 236    | 5     | AAO17051 Human alb |
| 9          | 32    | 100.0       | 241    | 5     | AAO16984 Alpha-MSH |
| 10         | 32    | 100.0       | 241    | 7     | ABU10025 Alpha-MSH |
| 11         | 32    | 100.0       | 242    | 5     | AAO16985 Alpha-MSH |
| 12         | 32    | 100.0       | 244    | 5     | AAO16986 Alpha-MSH |
| 13         | 32    | 100.0       | 245    | 5     | AAO16987 Alpha-MSH |
| 14         | 32    | 100.0       | 245    | 5     | AAO16988 Alpha-MSH |
| 15         | 32    | 100.0       | 268    | 5     | AAO16989 Alpha-MSH |
| 16         | 32    | 100.0       | 268    | 7     | ABU10026 Human ser |
| 17         | 32    | 100.0       | 303    | 2     | AAr14178 Human ser |
| 18         | 32    | 100.0       | 327    | 4     | AAU32564 Novel hum |
| 19         | 32    | 100.0       | 327    | 4     | AAU29942 Novel hum |
| 20         | 32    | 100.0       | 327    | 4     | AAU32995 Novel hum |
| 21         | 32    | 100.0       | 327    | 4     | AAU33287 Novel hum |
| 22         | 32    | 100.0       | 373    | 1     | AAp90387 N-termina |
| 23         | 32    | 100.0       | 388    | 1     | AAp90389 N-termina |
| 24         | 32    | 100.0       | 389    | 1     | AAp90390 N-termina |
| 25         | 32    | 100.0       | 390    | 1     | AAp90391 N-termina |

|    |    |       |     |   |                    |
|----|----|-------|-----|---|--------------------|
| 26 | 32 | 100.0 | 401 | 4 | AAU29876 Novel hum |
| 27 | 32 | 100.0 | 407 | 1 | AAp90392 N-termina |
| 28 | 32 | 100.0 | 463 | 2 | AAr14179 Human ser |
| 29 | 32 | 100.0 | 500 | 7 | ADD32019 Heterolog |
| 30 | 32 | 100.0 | 550 | 4 | AAU29877 Novel hum |
| 31 | 32 | 100.0 | 584 | 6 | ABG72381 Mature hu |
| 32 | 32 | 100.0 | 585 | 1 | AAp93344 Sequence  |
| 33 | 32 | 100.0 | 585 | 1 | AAp90388 Mature hu |
| 34 | 32 | 100.0 | 585 | 1 | AAp91422 Human nor |
| 35 | 32 | 100.0 | 585 | 1 | AAr05318 Human ser |
| 36 | 32 | 100.0 | 585 | 2 | AAr05457 Human ser |
| 37 | 32 | 100.0 | 585 | 2 | AAr26207 Human ser |
| 38 | 32 | 100.0 | 585 | 2 | AAr26362 Synthetic |
| 39 | 32 | 100.0 | 585 | 2 | AAr20029 Human ser |
| 40 | 32 | 100.0 | 585 | 2 | AAr80301 Human ser |
| 41 | 32 | 100.0 | 585 | 2 | AAO20111 HSA prote |
| 42 | 32 | 100.0 | 585 | 2 | AAW59841 Mature pr |
| 43 | 32 | 100.0 | 585 | 3 | AAy84873 Amino aci |
| 44 | 32 | 100.0 | 585 | 3 | AAy83946 Yeast cod |
| 45 | 32 | 100.0 | 585 | 4 | AAW52567 Mature hu |

ALIGNMENTS

RESULT 1

AAO00108  
ID AAO00108 standard; protein; 116 AA.

AC AAO00108;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 14000.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

DR N-PSDB; AAI50039.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 14000; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 116 AA;

Query Match 100.0%; Score 32; DB 4; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 24; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

QY 1 QAADKAA 7  
 DB 11 QAADKAA 17  
 |||||

RESULT 2  
 AA002571  
 ID AA002571 standard; protein; 143 AA.  
 XX  
 AC AA002571;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 16463.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US004927.  
 XX  
 PR 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 PI WPI; 2001-514838/56.  
 DR N-PSDB; AAI82502.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX  
 PS Claim 20; SEQ ID NO 16463; 1399pp + Sequence Listing; English.  
 CC  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 143 AA;

Query Match 100.0%; Score 32; DB 4; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 DB 11 QAADKAA 17  
 |||||

RESULT 3  
 AAU29875  
 ID AAU29875 standard; protein; 192 AA.  
 XX  
 AC AAU29875;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Novel human secreted protein #366.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179449-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 16-APR-2001; 2001WO-US008656.  
 XX  
 PR 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 PI WPI; 2001-611725/70.  
 DR  
 XX  
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX  
 PS Claim 20; Page 206; 765pp; English.  
 CC  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 192 AA;

Query Match 100.0%; Score 32; DB 4; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 DB 176 QAADKAA 182  
 |||||

RESULT 4  
 AA017048  
 ID AA017048 standard; protein; 195 AA.

XX AA017048;  
AC  
XX  
XX  
DT 29-MAY-2002 (first entry)  
XX  
XX Human serum albumin (1-195) SEQ ID NO: 56.  
DE  
DE  
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX  
XX Homo sapiens.  
OS  
XX WO200206316-A2.  
XX  
XX 24-JAN-2002.  
XX  
XX 16-JUL-2001; 2001WO-US022263.  
XX  
XX 14-JUL-2000; 2000US-0218381P.  
XX  
XX 18-AUG-2000; 2000US-0226382P.  
XX  
XX 06-OCT-2000; 2000US-0238380P.  
XX  
XX 29-DEC-2000; 2000US-0258764P.  
XX  
XX 14-JUN-2001; 2001US-0298317P.  
XX  
XX (ZVCO-) ZVCO INC.  
FA  
XX  
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
PI WPI; 2002-195801/25.  
XX  
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory  
PT or autoimmune disorders.  
XX  
XX Example 2; Page 46; 89pp; English.  
XX  
XX The present invention relates to a nucleic acid comprising a sequence  
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
CC hormone (MSH) concatamer. The sequences are useful for treating an  
CC individual suffering from, or at risk of, a disorder of the immune system  
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
XX  
XX Sequence 195 AA;  
SQ  
Query Match 100.0%; Score 32; DB 5; Length 195;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
Db 170 QAADKAA 176  
RESULT 5  
ABU10022  
ID ABU10022 standard; protein; 195 AA.  
XX  
XX ABU10022;  
AC  
XX 31-JUL-2003 (first entry)  
DT  
XX Human serum albumin residues 1-195.  
DE  
XX

KW Bladder disorder; cytostatic; antiinflammatory; immune response;  
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
KW bladder cancer; tumour; interstitial cystitis; inflammation;  
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;  
KW serum albumin.  
XX  
XX Homo sapiens.  
OS  
XX US2002193332-A1.  
PN  
XX 19-DEC-2002.  
PD  
XX 12-FEB-2002; 2002US-00074956.  
XX  
XX 12-FEB-2001; 2001US-0268175P.  
XX  
XX (HEDL/) HEDLEY M L.  
FA  
XX Hedley ML;  
PI  
XX WPI; 2003-447327/42.  
XX  
XX Modulating immune responses in a mammal with a bladder disorder e.g.  
PT bladder cancer, by administering nucleic acids comprising un-methylated  
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
PT the mammal.  
XX  
XX Example 2; Page 9; 17pp; English.  
XX  
XX The invention describes a method of modulating an immune response in a  
CC mammal, comprising identifying a mammal that has or is at risk for having  
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated  
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;  
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
CC method is useful for modulating immune response in a mammal having a  
CC bladder disorder, where administration of (N1) results in an amelioration  
CC of one or more symptoms of the disorder. Preferably, the method is useful  
CC for modulating immune response in a mammal having bladder cancer (where  
CC administration of (N1) results in a decrease in tumour size or activity),  
CC or for modulating immune response in a mammal having interstitial  
CC cystitis (where administration of (N1) results in a modulation of the  
CC immune response from Th2 response to a Th1 response). The method is also  
CC useful for modulating immune response in a mammal having bladder disorder  
CC that is characterised by inflammation which is associated with symptoms  
CC of interstitial cystitis or associated with a disruption of the integrity  
CC of the bladder lining. This is the amino acid sequence of human serum  
CC albumin residues 1-195 that can be used in the creation of melanocyte  
CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the  
CC fusion protein when expressed in mammalian cells  
XX  
XX Sequence 195 AA;  
SQ  
Query Match 100.0%; Score 32; DB 7; Length 195;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
Db 170 QAADKAA 176  
RESULT 6  
AAY83947  
ID AAY83947 standard; protein; 204 AA.  
XX  
XX AAY83947;  
AC  
XX 28-JUL-2000 (first entry)  
DT  
XX Yeast codon-biased recombinant HSA protein fragment HSA-I.  
DE  
XX Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW



overlapping oligonucleotide; expression vector.

OS Homo sapiens.  
OS Synthetic.  
XX CN1239103-A.  
XX PD 22-DEC-1999.  
XX PF 17-JUN-1998; 98CN-00102506.  
XX PR 17-JUN-1998; 98CN-00102506.  
XX PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
XX PI Li S, Lu D;  
XX WPI; 2000-351198/31.  
XX DR N-PSDB; AAA10092.  
XX PT Process for preparing recombinant human serum albumin comprising yeast  
XX biased sex codons - uses a recombinant DNA technique.  
XX PS Example 1; Fig 3; 44pp; Chinese.  
XX CC The method relates to a method of recombinantly producing human serum  
XX albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
XX a yeast codon bias. The complete HSA gene (AAA10091) was generated as  
XX three synthetic fragments (AAA10092-A10094) joined by recombinant DNA  
XX technology. Each HSA fragment was synthesised from overlapping  
XX oligonucleotide fragments that were extended. This sequence represents  
XX the sequence of the HSA fragment HSA-I encoded by the human gene with a  
XX yeast codon bias. The invention also covers a recombinant expression  
XX vector, yeast host cells carrying the recombinant expression vector and  
XX the process for producing human serum albumin in the yeast host cell,  
XX especially in secretory mode  
XX SQ Sequence 204 AA;

Query Match 100.0%; Score 32; DB 3; Length 204;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
DB 177 QAADKAA 183  
|||||

RESULT 7  
AAU29874  
ID AAU29874 standard; protein; 214 AA.  
XX AC AAU29874;  
XX AC AAU29874;  
DT 18-DEC-2001 (first entry)  
XX DE Novel human secreted protein #365.  
XX KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX OS Homo sapiens.  
XX PN WO200179449-A2.  
XX PD 25-OCT-2001.  
XX PF 16-APR-2001; 2001WO-US008656.  
XX PR 18-APR-2000; 2000US-00552929.  
XX PR 26-JAN-2001; 2001US-00770160.

(HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.  
XX PS Claim 20; Page 205; 765pp; English.  
XX CC The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation, to regulate haematopoiesis, and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU3304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX SQ Sequence 214 AA;

Query Match 100.0%; Score 32; DB 4; Length 214;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
DB 195 QAADKAA 201  
|||||

RESULT 8  
AAO17051  
ID AAO17051 standard; protein; 236 AA.  
XX AC AAO17051;  
XX AC AAO17051;  
DT 29-MAY-2002 (first entry)  
XX DE Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.  
XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;  
KW antidiabetic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX OS Homo sapiens.  
XX PN WO200206316-A2.  
XX PD 24-JAN-2002.  
XX PF 16-JUL-2001; 2001WO-US022263.  
XX PR 14-JUL-2000; 2000US-0218381P.  
XX PR 18-AUG-2000; 2000US-0226382P.  
XX PR 06-OCT-2000; 2000US-0238380P.  
XX PR 29-DEC-2000; 2000US-0258764P.  
XX PR 14-JUN-2001; 2001US-0298317P.  
XX PA (ZYCO-) ZYCOS INC.

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.  
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory  
PT or autoimmune disorders.  
XX Example 2; Page 48; 89pp; English.  
XX The present invention relates to a nucleic acid comprising a sequence  
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
CC hormone (MSH) concatamer. The sequences are useful for treating an  
CC individual suffering from, or at risk of, a disorder of the immune system  
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a peptide described in the exemplification of the  
CC invention  
XX Sequence 236 AA;  
SQ Query Match 100.0%; Score 32; DB 5; Length 236;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
DB 194 QAADKAA 200  
RESULT 9  
AAO16984  
ID AAO16984 standard; protein; 241 AA.  
XX AAO16984;  
XX 29-MAY-2002 (first entry)  
XX Alpha-MSH construct protein fragment SEQ ID NO: 59.  
DE Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW Alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;  
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX Unidentified.  
XX WO200206316-A2.  
XX 24-JAN-2002.  
XX 16-JUL-2001; 2001WO-US022263.  
XX 14-JUL-2000; 2000US-0218381P.  
XX 18-AUG-2000; 2000US-0226382P.  
XX 06-OCT-2000; 2000US-023380P.  
XX 29-DEC-2000; 2000US-0258764P.  
XX 14-JUN-2001; 2001US-0298317P.  
XX (ZYCO-) ZYCOS INC.  
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.  
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory

PT or autoimmune disorders.  
XX Example 2; Page 4-5; 89pp; English.  
XX The present invention relates to a nucleic acid comprising a sequence  
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
CC hormone (MSH) concatamer. The sequences are useful for treating an  
CC individual suffering from, or at risk of, a disorder of the immune system  
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
XX Sequence 241 AA;  
SQ Query Match 100.0%; Score 32; DB 5; Length 241;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
DB 194 QAADKAA 200  
RESULT 10  
ABU10025  
ID ABU10025 standard; protein; 241 AA.  
XX ABU10025;  
XX 31-JUL-2003 (first entry)  
XX Alpha-MSH/Serum albumin fusion protein H9.  
DE Bladder disorder; cytostatic; anti-inflammatory; immune response;  
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
KW bladder cancer; tumour; interstitial cystitis; inflammation;  
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;  
KW serum albumin; fusion protein; H9.  
XX Homo sapiens.  
OS Synthetic.  
OS US2002193332-A1.  
XX 13-DEC-2002.  
XX 12-FEB-2002; 2002US-00074956.  
XX 12-FEB-2001; 2001US-0268175P.  
XX (HEDL/) HEDLEY M L.  
XX Hedley ML;  
XX WPI; 2003-447327/42.  
XX Modulating immune responses in a mammal with a bladder disorder e.g.  
PT bladder cancer, by administering nucleic acids comprising un-methylated  
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
PT the mammal.  
XX Example 2; Page 9; 17pp; English.  
XX The invention describes a method of modulating an immune response in a  
CC mammal, comprising identifying a mammal that has or is at risk for having  
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated  
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;  
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
CC method is useful for modulating immune response in a mammal having a

CC bladder disorder, where administration of (N1) results in an amelioration  
CC of one or more symptoms of the disorder. Preferably, the method is useful  
CC for modulating immune response in a mammal having bladder cancer (where  
CC administration of (N1) results in a decrease in tumour size or activity),  
CC or for modulating immune response in a mammal having interstitial  
CC cystitis (where administration of (N1) results in a modulation of the  
CC immune response from Th2 response to a Th1 response). The method is also  
CC useful for modulating immune response in a mammal having bladder disorder  
CC that is characterised by inflammation which is associated with symptoms  
CC of interstitial cystitis or associated with a disruption of the integrity  
CC of the bladder lining. This is the amino acid sequence of a human serum  
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the  
CC treatment of bladder disorders  
XX  
SQ Sequence 241 AA;

Query Match 100.0%; Score 32; DB 7; Length 241;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAAKAA 7  
| | | | |  
DB 194 QAAAKAA 200

RESULT 11  
AAO16985  
ID AAO16985 standard; protein; 242 AA.

XX  
AC AAO16985;  
XX  
DT 29-MAY-2002 (first entry)  
XX  
DE Alpha-MSH construct protein fragment SEQ ID NO: 70.  
XX  
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;  
KW antiaesthetic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.

XX Unidentified.

XX WO200206316-A2.

XX 24-JAN-2002.

XX 16-JUL-2001; 2001WO-US022263.

XX 14-JUL-2000; 2000US-0218381P.

XX 18-AUG-2000; 2000US-0228382P.

XX 06-OCT-2000; 2000US-0238380P.

XX 29-DEC-2000; 2000US-0258764P.

XX 14-JUN-2001; 2001US-0298317P.

XX (ZYCO-) ZYCOS INC.

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.

XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
XX stimulating hormone concatamer or its analog, for treating inflammatory  
XX or autoimmune disorders.

XX Disclosure; Page 5; 89pp; English.

XX The present invention relates to a nucleic acid comprising a sequence  
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating  
XX hormone (MSH) concatamer. The sequences are useful for treating an  
XX individual suffering from, or at risk of, a disorder of the immune system.

CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
XX

SQ Sequence 242 AA;

Query Match 100.0%; Score 32; DB 5; Length 242;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAAKAA 7  
| | | | |  
DB 194 QAAAKAA 200

RESULT 12  
AAO16986  
ID AAO16986 standard; protein; 244 AA.

XX  
AC AAO16986;

XX  
DT 29-MAY-2002 (first entry)

XX  
DE Alpha-MSH construct protein fragment SEQ ID NO: 71.

XX  
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;  
KW antiaesthetic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.

XX Unidentified.

XX WO200206316-A2.

XX 24-JAN-2002.

XX 16-JUL-2001; 2001WO-US022263.

XX 14-JUL-2000; 2000US-0218381P.

XX 18-AUG-2000; 2000US-0228382P.

XX 06-OCT-2000; 2000US-0238380P.

XX 29-DEC-2000; 2000US-0258764P.

XX 14-JUN-2001; 2001US-0298317P.

XX (ZYCO-) ZYCOS INC.

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.

XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
XX stimulating hormone concatamer or its analog, for treating inflammatory  
XX or autoimmune disorders.

XX Disclosure; Page 5; 89pp; English.

XX The present invention relates to a nucleic acid comprising a sequence  
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating  
XX hormone (MSH) concatamer. The sequences are useful for treating an  
XX individual suffering from, or at risk of, a disorder of the immune system  
XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
XX present sequence is a protein described in the exemplification of the  
XX invention

SQ Sequence 244 AA;  
 Query Match 100.0%; Score 32; DB 5; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 |||||  
 Db 194 QAADKAA 200

RESULT 13  
 AAO16987  
 ID AAO16987 standard; protein; 245 AA.  
 XX AC AAO16987;  
 XX DT 29-MAY-2002 (first entry)  
 XX DE Alpha-MSH construct protein fragment SEQ ID NO: 72.  
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
 KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
 KW diabetes; uveitis; coeliac disease.  
 XX OS Unidentified.  
 XX PN WO200206316-A2.  
 XX PD 24-JAN-2002.  
 XX PF 16-JUL-2001; 2001WO-US022263.  
 XX PR 14-JUL-2000; 2000US-0218381P.  
 XX PR 18-AUG-2000; 2000US-0226382P.  
 XX PR 06-OCT-2000; 2000US-0238380P.  
 XX PR 29-DEC-2000; 2000US-0258764P.  
 XX PR 14-JUN-2001; 2001US-0298317P.  
 XX PA (ZYCO-) ZYCOS INC.  
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 XX WPI; 2002-195801/25.  
 XX DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 PT stimulating hormone concatamer or its analog, for treating inflammatory  
 PT or autoimmune disorders.  
 XX PS Disclosure; Page 5; 89pp; English.  
 XX CC The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 CC hormone (MSH) concatamer. The sequences are useful for treating an  
 CC individual suffering from, or at risk of, a disorder of the immune system  
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 CC present sequence is a protein described in the exemplification of the  
 CC invention  
 XX SQ Sequence 245 AA;  
 Query Match 100.0%; Score 32; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 |||||  
 Db 194 QAADKAA 200

RESULT 15  
 AAO16989  
 ID AAO16989 standard; protein; 268 AA.

Db 194 QAADKAA 200  
 |||||  
 RESULT 14  
 AAO16988  
 ID AAO16988 standard; protein; 245 AA.  
 XX AC AAO16988;  
 XX DT 29-MAY-2002 (first entry)  
 XX DE Alpha-MSH construct protein fragment SEQ ID NO: 73.  
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
 KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
 KW diabetes; uveitis; coeliac disease.  
 XX OS Unidentified.  
 XX PN WO200206316-A2.  
 XX PD 24-JAN-2002.  
 XX PF 16-JUL-2001; 2001WO-US022263.  
 XX PR 14-JUL-2000; 2000US-0218381P.  
 XX PR 18-AUG-2000; 2000US-0226382P.  
 XX PR 06-OCT-2000; 2000US-0238380P.  
 XX PR 29-DEC-2000; 2000US-0258764P.  
 XX PR 14-JUN-2001; 2001US-0298317P.  
 XX PA (ZYCO-) ZYCOS INC.  
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 XX WPI; 2002-195801/25.  
 XX DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 PT stimulating hormone concatamer or its analog, for treating inflammatory  
 PT or autoimmune disorders.  
 XX PS Disclosure; Page 5; 89pp; English.  
 XX CC The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 CC hormone (MSH) concatamer. The sequences are useful for treating an  
 CC individual suffering from, or at risk of, a disorder of the immune system  
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 CC present sequence is a protein described in the exemplification of the  
 CC invention  
 XX SQ Sequence 245 AA;  
 Query Match 100.0%; Score 32; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
 |||||  
 Db 194 QAADKAA 200

RESULT 15  
 AAO16989  
 ID AAO16989 standard; protein; 268 AA.

XX AAO16989;  
AC  
XX  
XX 29-MAY-2002 (first entry)  
DT  
XX  
XX Alpha-MSH construct protein fragment SEQ ID NO: 60.  
DE  
XX  
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;  
KW antistaphylococcal; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX  
XX Unidentified.  
OS  
XX  
XX WO200206316-A2.  
FN  
XX  
XX 24-JAN-2002.  
PD  
XX  
XX 16-JUL-2001; 2001WO-US022263.  
PF  
XX  
XX 14-JUL-2000; 2000US-0218381P.  
PR  
XX 18-AUG-2000; 2000US-0226382P.  
PR  
XX 06-OCT-2000; 2000US-0238380P.  
PR  
XX 29-DEC-2000; 2000US-0258764P.  
PR  
XX 14-JUN-2001; 2001US-0298317P.  
PR  
XX (ZYCO-) ZYCOS INC.  
PA  
XX  
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
FI  
XX  
XX WPI; 2002-195801/25.  
DR  
XX  
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory  
PT or autoimmune disorders.  
PT  
XX  
XX Example 2; Page 5; 89pp; English.  
PS  
XX  
XX The present invention relates to a nucleic acid comprising a sequence  
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
CC hormone (MSH) concatamer. The sequences are useful for treating an  
CC individual suffering from, or at risk of, a disorder of the immune system  
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
CC  
XX  
XX Sequence 268 AA;  
SQ  
Query Match 100.0%; Score 32; DB 5; Length 268;  
Best Local Similarity 100.0%; Pred.No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 QAADKAA 7  
DB 194 QAADKAA 200  
Search completed: April 19, 2004, 11:51:18  
Job time : 5.11727 secs

| NO. | SCORE | PERCENT | GRADE | STUDENT                 |
|-----|-------|---------|-------|-------------------------|
| 1   | 32    | 100.0   | 104   | 12 US-10-424-599-269254 |
| 2   | 32    | 100.0   | 195   | 13 US-10-074-956-24     |
| 3   | 32    | 100.0   | 241   | 13 US-10-074-956-27     |
| 4   | 32    | 100.0   | 268   | 13 US-10-074-956-28     |
| 5   | 32    | 100.0   | 585   | 9 US-09-939-452-2       |
| 6   | 32    | 100.0   | 585   | 10 US-09-932-613-445    |
| 7   | 32    | 100.0   | 585   | 10 US-09-984-010-26     |
| 8   | 32    | 100.0   | 585   | 10 US-09-833-041-18     |
| 9   | 32    | 100.0   | 585   | 10 US-09-833-117-18     |
| 10  | 32    | 100.0   | 585   | 10 US-09-932-22-445     |
| 11  | 32    | 100.0   | 585   | 10 US-09-832-501-18     |
| 12  | 32    | 100.0   | 585   | 11 US-09-833-118-18     |
| 13  | 32    | 100.0   | 585   | 11 US-09-833-245-18     |
| 14  | 32    | 100.0   | 585   | 12 US-10-434-999-11     |
| 15  | 32    | 100.0   | 585   | 12 US-10-435-000-31     |
|     |       |         |       | Sequence 269254,        |
|     |       |         |       | Sequence 24, Appl       |
|     |       |         |       | Sequence 27, Appl       |
|     |       |         |       | Sequence 28, Appl       |
|     |       |         |       | Sequence 2, Appl        |
|     |       |         |       | Sequence 445, App       |
|     |       |         |       | Sequence 26, Appl       |
|     |       |         |       | Sequence 18, Appl       |
|     |       |         |       | Sequence 18, Appl       |
|     |       |         |       | Sequence 11, Appl       |
|     |       |         |       | Sequence 31, Appl       |

;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS  
;; FILE REFERENCE: 08191-022001  
;; CURRENT APPLICATION NUMBER: US/10/074,956  
;; PRIOR FILING DATE: 2002-06-10  
;; PRIOR APPLICATION NUMBER: 60/268,175  
;; PRIOR FILING DATE: 2001-02-12  
;; NUMBER OF SEQ ID NOS: 29  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 24  
;; LENGTH: 195  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-074-956-24

Query Match 100.0%; Score 32; DB 13; Length 195;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
Db 170 QAADKAA 176

RESULT 3  
US-10-074-956-27  
;; Sequence 27, Application US/10074956  
;; Publication No. US2002019332A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Hedley, Mary Lynne  
;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS  
;; FILE REFERENCE: 08191-022001  
;; CURRENT APPLICATION NUMBER: US/10/074,956  
;; CURRENT FILING DATE: 2002-06-10  
;; PRIOR APPLICATION NUMBER: 60/268,175  
;; PRIOR FILING DATE: 2001-02-12  
;; NUMBER OF SEQ ID NOS: 29  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 27  
;; LENGTH: 241  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-074-956-27

Query Match 100.0%; Score 32; DB 13; Length 241;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 4  
US-10-074-956-28  
;; Sequence 28, Application US/10074956  
;; Publication No. US2002019332A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Hedley, Mary Lynne  
;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS  
;; FILE REFERENCE: 08191-022001  
;; CURRENT APPLICATION NUMBER: US/10/074,956  
;; CURRENT FILING DATE: 2002-06-10  
;; PRIOR APPLICATION NUMBER: 60/268,175  
;; PRIOR FILING DATE: 2001-02-12  
;; NUMBER OF SEQ ID NOS: 29  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 28  
;; LENGTH: 268  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-074-956-28

Query Match 100.0%; Score 32; DB 13; Length 268;

Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 5  
US-09-929-552-2  
;; Sequence 2, Application US/09929552  
;; Patent No. US20020123080A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sonnenschein, Carlos  
;; Soto, Ana M.  
;; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
;; NUMBER OF SEQUENCES: 2  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Medlen & Carroll, LLP  
;; STREET: 220 Montgomery Street, Suite 2200  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: United States of America  
;; ZIP: 94104  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/929,552  
;; FILING DATE: 14-Aug-2001  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/769,746  
;; FILING DATE: 19-DEC-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Carroll, Peter G.  
;; REGISTRATION NUMBER: 32,837  
;; REFERENCE/DOCKET NUMBER: MBRI-02584  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 705-8410  
;; TELEFAX: (415) 397-8338  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 585 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-929-552-2

Query Match 100.0%; Score 32; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7  
Db 170 QAADKAA 176

RESULT 6  
US-09-932-613-445  
;; Sequence 445, Application US/09932613  
;; Publication No. US20030091565A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Human Genome Sciences, Inc.  
;; APPLICANT: Beltzer, James P.  
;; APPLICANT: Potter, M. Daniel  
;; APPLICANT: Fleming, Tony J.  
;; APPLICANT: Rosen, Craig A.  
;; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
;; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US

```
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QAADKAA 7
Db      170 QAADKAA 176

RESULT 7
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; APPLICANT: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; ADDRESS: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QAADKAA 7
Db      170 QAADKAA 176

RESULT 8
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS45
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18
```

```
Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 QAADKAA 7
Db      170 QAADKAA 176
```

```
RESULT 9
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS43
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18
```

```
Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 QAADKAA 7
Db      170 QAADKAA 176
```

```
RESULT 10
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
```



```

; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match 100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
   |||||
Db 170 QAADKAA 176

RESULT 11
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match 100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
   |||||
Db 170 QAADKAA 176

RESULT 12
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358

```

```

; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 32; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 15
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Camecon, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 32; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

Search completed: April 19, 2004, 12:54:58
Job time : 2.94737 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 2.14589 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_76\_89

Perfect score: 70

Sequence: 1 TVATLRITYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 70    | 100.0       | 585    | 1  | US-08-153-799-14  |
| 2          | 70    | 100.0       | 585    | 1  | US-08-448-196A-3  |
| 3          | 70    | 100.0       | 585    | 2  | US-08-984-176-1   |
| 4          | 70    | 100.0       | 585    | 2  | US-08-702-572-2   |
| 5          | 70    | 100.0       | 585    | 3  | US-08-769-746-2   |
| 6          | 70    | 100.0       | 585    | 4  | US-10-153-064-5   |
| 7          | 70    | 100.0       | 609    | 1  | US-08-222-619-3   |
| 8          | 70    | 100.0       | 609    | 1  | US-08-433-037-4   |
| 9          | 70    | 100.0       | 609    | 4  | US-08-897-956A-2  |
| 10         | 70    | 100.0       | 609    | 4  | US-10-153-064-7   |
| 11         | 70    | 100.0       | 609    | 4  | US-09-976-594-977 |
| 12         | 70    | 100.0       | 609    | 5  | PCT-US95-04075-3  |
| 13         | 70    | 100.0       | 610    | 2  | US-08-797-689-2   |
| 14         | 70    | 100.0       | 610    | 4  | US-09-984-186-2   |
| 15         | 70    | 100.0       | 651    | 4  | US-10-153-064-133 |
| 16         | 70    | 100.0       | 652    | 4  | US-10-153-064-96  |
| 17         | 70    | 100.0       | 652    | 4  | US-10-153-064-99  |
| 18         | 70    | 100.0       | 652    | 4  | US-10-153-064-105 |
| 19         | 70    | 100.0       | 652    | 4  | US-10-153-064-132 |
| 20         | 70    | 100.0       | 653    | 4  | US-10-153-064-131 |
| 21         | 70    | 100.0       | 656    | 4  | US-10-153-064-130 |
| 22         | 70    | 100.0       | 660    | 4  | US-10-153-064-90  |
| 23         | 70    | 100.0       | 660    | 4  | US-10-153-064-93  |
| 24         | 70    | 100.0       | 668    | 4  | US-10-153-064-102 |
| 25         | 70    | 100.0       | 676    | 4  | US-10-153-064-95  |
| 26         | 70    | 100.0       | 676    | 4  | US-10-153-064-98  |
| 27         | 70    | 100.0       | 676    | 4  | US-10-153-064-104 |

|    |    |       |      |   |                     |                   |
|----|----|-------|------|---|---------------------|-------------------|
| 28 | 70 | 100.0 | 676  | 4 | US-10-153-064-127   | Sequence 127, App |
| 29 | 70 | 100.0 | 676  | 4 | US-10-153-064-129   | Sequence 129, App |
| 30 | 70 | 100.0 | 677  | 4 | US-10-153-064-125   | Sequence 125, App |
| 31 | 70 | 100.0 | 680  | 4 | US-10-153-064-123   | Sequence 123, App |
| 32 | 70 | 100.0 | 684  | 4 | US-10-153-064-92    | Sequence 92, Appl |
| 33 | 70 | 100.0 | 692  | 4 | US-10-153-064-101   | Sequence 101, App |
| 34 | 70 | 100.0 | 783  | 1 | US-08-256-938-2     | Sequence 2, Appli |
| 35 | 70 | 100.0 | 787  | 1 | US-08-256-938-4     | Sequence 4, Appli |
| 36 | 70 | 100.0 | 787  | 2 | US-08-797-689-16    | Sequence 16, Appl |
| 37 | 70 | 100.0 | 787  | 4 | US-09-984-186-16    | Sequence 16, Appl |
| 38 | 70 | 100.0 | 978  | 4 | US-08-897-956A-3    | Sequence 3, Appli |
| 39 | 70 | 100.0 | 1184 | 4 | US-10-153-064-89    | Sequence 89, Appl |
| 40 | 62 | 88.6  | 583  | 1 | US-08-448-196A-6    | Sequence 6, Appli |
| 41 | 61 | 87.1  | 583  | 1 | US-08-448-196A-5    | Sequence 5, Appli |
| 42 | 58 | 82.9  | 582  | 1 | US-08-134-638-1     | Sequence 1, Appli |
| 43 | 58 | 82.9  | 583  | 1 | US-08-448-196A-4    | Sequence 4, Appli |
| 44 | 42 | 60.0  | 584  | 1 | US-08-448-196A-7    | Sequence 7, Appli |
| 45 | 40 | 57.1  | 104  | 4 | US-09-134-000C-5225 | Sequence 5225, Ap |

ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-n)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"

US-08-153-799-14  
Query Match 100.0%; Score 70; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 5.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 76 TVATLRETYGEMAD 89

## RESULT 2

US-08-448-196A-3  
Sequence 3, Application US/08448196A  
Patent No. 5780594

GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C.  
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
TITLE OF INVENTION: RELATED PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA  
STREET: MARSHALL SPACE FLIGHT CENTER  
CITY: HUNTSVILLE  
STATE: ALABAMA  
COUNTRY: USA  
ZIP: 35812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,196A  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: BROAD JR., ROBERT L.  
REGISTRATION NUMBER: 18,757  
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 205-544-0021  
TELEFAX: 205-544-0258  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 100.0%; Score 70; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 5.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 76 TVATLRETYGEMAD 89

## RESULT 3

US-08-984-176-1  
Sequence 1, Application US/08984176  
Patent No. 5948609

GENERAL INFORMATION:  
APPLICANT: CARTER, DANIEL C  
APPLICANT: HO, JOSEPH X  
APPLICANT: RUKER, FLORIAN  
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER  
FILE REFERENCE: 08/984,176  
CURRENT APPLICATION NUMBER: US/08/984,176  
CURRENT FILING DATE: 1997-12-03  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 1  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-984-176-1

Query Match 100.0%; Score 70; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 5.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 76 TVATLRETYGEMAD 89

## RESULT 4

US-08-702-572-2  
Sequence 2, Application US/08702572  
Patent No. 5965386

GENERAL INFORMATION:  
APPLICANT: Kerry-Williams, Sean M  
APPLICANT: Gilbert, Sarah C  
TITLE OF INVENTION: Yeast Strains and Modified Albumins  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Centeon L.L.C.  
STREET: 1020 First Avenue  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406-1310

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,572  
FILING DATE: 11-NOV-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 95/23857  
FILING DATE: 1-MAR-1995  
APPLICATION NUMBER: GB 9404270.2  
FILING DATE: 5-MAR-1994  
ATTORNEY/AGENT INFORMATION:

NAME: Naomi Biswas  
REGISTRATION NUMBER: 38,384  
REFERENCE/DOCKET NUMBER: CE0114 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610/878/4294  
TELEFAX: 610/878/4221  
INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 585 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-702-572-2
Query Match      100.0%; Score 70; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 585 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-769-746-2
Query Match      100.0%; Score 70; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 50/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match      100.0%; Score 70; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 609 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
US-08-222-619-3
Query Match      100.0%; Score 70; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      100 TVATLRETYGEMAD 113

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```

APPLICANT: Barr, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschopp, Juerg F.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 91082  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 70; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 9  
US-08-897-956A-2  
Sequence 2, Application US/08897956A  
Patent No. 6423512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 70; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 10  
US-10-153-064-7  
Sequence 7, Application US/10153064  
Patent No. 6663485  
GENERAL INFORMATION:  
APPLICANT: Bell et al.  
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
FILE REFERENCE: PF556  
CURRENT APPLICATION NUMBER: US/10/153,064  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 60/293,212  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 7  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 70; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 11  
US-09-976-594-977  
Sequence 977, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 977  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1  
US-09-976-594-977

Query Match 100.0%; Score 70; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 12  
PCT-US95-04075-3  
Sequence 3, Application PC/TUS9504075  
GENERAL INFORMATION:  
APPLICANT: AMGEN INC.

;; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
;; TITLE OF INVENTION: Protein  
;; NUMBER OF SEQUENCES: 33  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Amgen Center, Patent Operations/RRC  
;; STREET: 1840 DeHavilland Drive  
;; CITY: Thousand Oaks  
;; STATE: California  
;; COUNTRY: U.S.  
;; ZIP: 91320-1789  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/04075  
;; FILING DATE:  
;; CLASSIFICATION:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 609 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 100.0%; Score 70; DB 5; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 13  
US-08-797-689-2  
; Sequence 2, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittion, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (Patentin)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38,619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-797-689-2

Query Match 100.0%; Score 70; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
Db 100 TVATLRETYGEMAD 113

RESULT 14  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. 6686179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittion, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (Patentin)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:

```
/
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 610 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/   MOLECULE TYPE: protein
/   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 70; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      100 TVATLRETYGEMAD 113

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
;   APPLICANT: Bell et al.
;   TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
;   FILE REFERENCE: PF356
;   CURRENT APPLICATION NUMBER: US/10/153,064
;   CURRENT FILING DATE: 2002-05-24
;   PRIOR APPLICATION NUMBER: 60/293,212
;   PRIOR FILING DATE: 2001-05-25
;   NUMBER OF SEQ ID NOS: 137
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 133
;   LENGTH: 651
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 70; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      142 TVATLRETYGEMAD 155

Search completed: April 19, 2004, 12:05:19
Job time : 2.14589 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.09695 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: Pirl.\*

2: Pirl2.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description           |
|------------|-------|-------------|--------|----------|-----------------------|
| 1          | 47    | 100.0       | 600    | 2 A47391 | serum albumin prec    |
| 2          | 47    | 100.0       | 608    | 1 ABRTS  | serum albumin prec    |
| 3          | 47    | 100.0       | 609    | 1 ABHUS  | serum albumin prec    |
| 4          | 47    | 100.0       | 609    | 2 JC5838 | albumin - Mongolia    |
| 5          | 43    | 91.5        | 453    | 2 A05139 | serum albumin - mo    |
| 6          | 43    | 91.5        | 607    | 1 ABBOS  | serum albumin prec    |
| 7          | 43    | 91.5        | 607    | 1 ABBOS  | serum albumin prec    |
| 8          | 43    | 91.5        | 607    | 1 ABBOS  | serum albumin prec    |
| 9          | 40    | 85.1        | 605    | 1 ABFGS  | serum albumin prec    |
| 10         | 39    | 83.0        | 608    | 2 S57632 | serum albumin prec    |
| 11         | 35    | 74.5        | 614    | 2 S59517 | serum albumin prec    |
| 12         | 35    | 74.5        | 615    | 1 ABBOS  | serum albumin prec    |
| 13         | 35    | 74.5        | 884    | 2 G36787 | serum albumin prec    |
| 14         | 35    | 74.5        | 1668   | 2 A02722 | protein T4012.9 [i    |
| 15         | 34    | 72.3        | 377    | 1 B0AG38 | IgA-specific metal    |
| 16         | 34    | 72.3        | 377    | 2 AF3249 | virB10 protein - A    |
| 17         | 34    | 72.3        | 680    | 2 A43800 | component of type     |
| 18         | 34    | 72.3        | 787    | 2 A48819 | nuclear autoantige    |
| 19         | 33    | 70.2        | 245    | 2 C55206 | nuclear autoantige    |
| 20         | 33    | 70.2        | 245    | 2 C51243 | thiamin biosynthes    |
| 21         | 33    | 70.2        | 245    | 2 A36091 | thiamin biosynthes    |
| 22         | 33    | 70.2        | 343    | 2 H84607 | hypothetical prote    |
| 23         | 33    | 70.2        | 399    | 2 H84108 | ABC transporter BH    |
| 24         | 33    | 70.2        | 407    | 2 G83477 | hypothetical sigma-70 |
| 25         | 33    | 70.2        | 433    | 2 B82965 | hypothetical prote    |
| 26         | 33    | 70.2        | 489    | 2 F70401 | Flagellar M-ring p    |
| 27         | 33    | 70.2        | 487    | 1 A3XL72 | 74K albumin precur    |
| 28         | 33    | 70.2        | 713    | 2 AC3220 | catalase C [impor     |
| 29         | 33    | 70.2        | 760    | 2 JQ1383 | hypothetical 84K p    |

30 33 70.2 906 2 G96621 probable disease r  
31 33 1364 2 T40839 hypothetical prote  
32 32 68.1 106 2 T00722 hypothetical prote  
33 32 68.1 175 2 T02557 hypothetical prote  
34 32 68.1 207 2 T10285 hypothetical prote  
35 32 68.1 412 2 F91008 probable oxidoredu  
36 32 68.1 412 2 H85852 probable oxidoredu  
37 32 68.1 486 2 G94433 hypothetical prote  
38 32 68.1 559 2 T12680 peroxisomal target  
39 32 68.1 721 2 D82934 ATP-dependent zinc  
40 32 68.1 732 2 S23001 tral protein - Esc  
41 32 68.1 1032 2 G89427 protein T08D2.3 [i  
42 32 68.1 1156 2 T43326 germline RNA helic  
43 32 68.1 1172 2 T32759 hypothetical prote  
44 32 68.1 1291 2 T00019 period protein hom  
45 32 68.1 3051 2 S42373 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A47391  
serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote  
A:Accession: A47391  
A>Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:21-194/Domain: serum albumin repeat homology <SA1>  
F:213-386/Domain: serum albumin repeat homology <SA2>  
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 47; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

DB 108 AKQEPERNE 116

##### RESULT 2

ABRTS

serum albumin precursor - rat

N:alternate names: preproalbumin

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-May-1979 #sequence\_revision 31-May-1979 #text\_change 22-Jun-1999

C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233

R:Sargent, T.D.; Yang, M.; Bonner, J.

Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981

A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.

A:Reference number: A93872; MUID:81223722; PMID:7017712

A:Accession: A93872

A:Molecule type: mRNA

A:Residues: 1-608 <SAR>

A:Cross-references: GB:V01222; GB:J00639; NID:g55627; PIDN:CAA24532.1; PID:g55628

R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.

J. Biol. Chem. 252, 6846-6855, 1977

A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analy:

A:Reference number: A92211; MUID:77249657; PMID:893447

A>Note: cleavages during protein maturation

A:Accession: A92211  
A:Molecule type: Protein  
A:Residues: 1-38 <STR>  
R:Risemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage  
A:Reference number: A91946; MUID:78109429; PMID:564345  
A:Accession: A91946  
A:Molecule type: Protein  
A:Residues: 25-222 <IS1>  
R:Risemura, S.; Ikenaka, T.  
J. Biochem. 79, 1183-1196, 1976  
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequence of the N-terminal fragment  
A:Reference number: A91940; MUID:76260153; PMID:956149  
A:Accession: A91940  
A:Molecule type: Protein  
A:Residues: 223-288/572-608 <IS2>  
A:Note: 262-Leu was also found  
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A:Reference number: A90758; MUID:79001617; PMID:80265  
A:Contents: annotation; copper binding  
R:Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: C45800  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <AR>  
R:Heard, J.  
Mol. Cell. Biol. 7, 2425-2434, 1987  
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved method  
A:Reference number: 157621; MUID:87286876; PMID:3475566  
A:Accession: 157621  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-5 <RES>  
A:Cross-references: GB:M16825; NID:9202828; PIDN:AAA40712.1; PID:9554412  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-608/Product: serum albumin #status experimental <WAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status experimental  
F:77-86, 99-115, 114-125, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 340-385, 384-393, 400-408, 413-425, 438-448, 461-475, 488-500, 513-520, 533-540, 553-560, 573-580, 596-608, 621-630, 643-650, 663-670, 683-690, 703-710, 723-730, 743-750, 763-770, 783-790, 803-810, 823-830, 843-850, 863-870, 883-890, 903-910, 923-930, 943-950, 963-970, 983-990, 1000-1008  
Query Match 100.0%; Score 47; DB 1; Length 608;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
|||||||  
Db 116 AKQEPERNE 124

RESULT 3  
ABUS  
serum albumin precursor [validated] - human  
N:Alternate names: preproalbumin  
N:Contains: kinetensin  
C:Species: Homo sapiens (man)  
C>Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000  
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36  
R:Liaw, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli  
A:Reference number: A93743; MUID:82081882; PMID:6171778  
A:Accession: A93743

A:Molecule type: mRNA  
A:Residues: 1-419, 'K', 421-609 <LAW>  
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:928591; PIDN:CAA2.  
R:Dugalczyk, A.; Law, S.W.; Dennison, C.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A:Reference number: A93936; MUID:82105994; PMID:6275391  
A:Accession: A93936  
A:Molecule type: mRNA  
A:Residues: 1-120, 'G', 122-609 <DUG>  
A:Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590  
R:Uranio, Y.; Watanabe, K.; Sakai, M.; Tamaki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A:Reference number: I39427; MUID:86140099; PMID:2419329  
A:Accession: I39427  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <URA>  
A:Cross-references: GB:M13075; NID:9178330; PIDN:AAA51688.1; PID:9553173  
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family  
A:Reference number: I59286; MUID:94181575; PMID:8134387  
A:Accession: I59286  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 282-290, 'KSRFDLQ' <WAT>  
A:Cross-references: GB:S69192; NID:9546032; PIDN:AA30282.1; PID:9546033  
A:Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl  
A:Reference number: I59313; MUID:94294404; PMID:8022807  
A:Accession: I59313  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 589-590, 'ALPREVKLLQVLP' <MAD>  
A:Cross-references: GB:S70799; NID:9547231; PIDN:AA311177.1; PID:9547232  
A:Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
R:Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: G08292  
A:Accession: G01747  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-120, 'G', 122-455 <MEN>  
A:Cross-references: EMBL:U22961; NID:9763428; PIDN:AAA64922.1; PID:9763431  
R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex  
A:Reference number: S55314; MUID:95275251; PMID:7755581  
A:Accession: S55314  
A:Molecule type: protein  
A:Residues: 19-27 <LED>  
R:Meloun, B.; Moravek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A:Title: Complete amino acid sequence of human serum albumin.  
A:Reference number: A91420; MUID:76187907; PMID:1225573  
A:Accession: A91420  
A:Molecule type: protein  
A:Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395  
R:Roehr, U.; Spittler, G.; Trippier, D.  
Justus Liebig Ann. Chem. 9, 881-884, 1988  
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A:Reference number: S06422  
A:Note: this paper is in German, with an English abstract  
A:Accession: S06422  
A:Molecule type: protein  
A:Residues: 25-48 <ROE>  
R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A:Title: Mass spectrometric identification of modifications to human serum albumin tree

A:Reference number: S36882; MUID:93384321; PMID:8373198  
A:Accession: S36882  
A:Molecule type: protein  
A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>  
R:Kausler, E.; Spiteley, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A:Title: Bruchstücke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmoleküle  
A:Reference number: S14999; MUID:92126241; PMID:11772598  
A:Accession: S17599  
A:Molecule type: protein  
A:Residues: 25-54;354-357;431-447 <KAU>  
A:Note: 49-Leu was also found  
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A:Reference number: A45800; MUID:93341406; PMID:2474609  
A:Accession: A45800  
A:Molecule type: protein  
A:Residues: 166-173 <CAR>  
R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Watkins, S.  
Biochem. Biophys. Res. Commun. 135, 983-988, 1986  
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-treated porcine stomach  
A:Reference number: A03239; MUID:86242180; PMID:3087352  
A:Accession: A03239  
A:Molecule type: protein  
A:Residues: 166-173, 'L' <MOG>  
R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A:Title: Mutations in genetic variants of human serum albumin found in Italy.  
A:Reference number: A38255; MUID:91062352; PMID:2247440  
A:Accession: C38255  
A:Molecule type: protein  
A:Residues: 76-111 <GAL1>  
A:Accession: B38255  
A:Molecule type: protein  
A:Residues: 82-105, 'K', 107-110 <GAL2>  
A:Note: this variant is designated albumin Vibo Valentia  
A:Accession: A38255  
A:Molecule type: protein  
A:Residues: 76-83, 'K', 85-106 <GAL3>  
A:Note: this variant is designated albumin Torino  
R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A:Title: The structural characterization and bilirubin-binding properties of albumin Her  
A:Reference number: S33298; MUID:9252504; PMID:8513793  
A:Accession: S33298  
A:Molecule type: protein  
A:Residues: 255-263, 'E', 265-281 <MIN1>  
A:Note: this variant is designated albumin Herborn  
R:Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta, F.  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A:Title: Two albumins with identical electrophoretic mobility are produced by different genes  
A:Reference number: S21078; MUID:92190239; PMID:1347703  
A:Accession: S21078  
A:Molecule type: protein  
A:Residues: 354-356, 'K', 358-378 <MIN2>  
A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported, Nature 358, 209-215, 1992  
R:He, X.M.; Carter, D.C.  
A:Title: Atomic structure and chemistry of human serum albumin.  
A:Reference number: A46756; MUID:92334427; PMID:1630489  
A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
R:Brown, J.R.; Shockley, P.; Behrens, P.Q.  
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
A:Reference number: A94442  
A:Contents: annotation; three-dimensional structure and disulfide bonds  
R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A:Title: Disulfide bonds in human serum albumin.  
A:Reference number: A9030  
A:Contents: annotation; disulfide bonds  
R:Jacobsen, C.  
Biochem. J. 171, 453-459, 1978

A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A:Reference number: A90299; MUID:78186630; PMID:656055  
A:Contents: annotation; bilirubin-binding site  
R:Peters, T.; Reed, R.G.  
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20  
A:Title: Serum albumin: conformation and active sites.  
A:Reference number: A34408  
A:Contents: annotation; binding sites  
R:Harper, M.E.; Dugalczyk, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene  
A:Reference number: A90028; MUID:83279982; PMID:6192711  
A:Contents: annotation; gene position  
R:Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid  
A:Reference number: A46755; MUID:76257808; PMID:955075  
A:Contents: annotation  
A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid) to lysine-199 of human serum albumin  
R:Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 265-268, 1992  
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate  
A:Reference number: A56294; MUID:92183881; PMID:1544460  
A:Contents: annotation  
A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in 1  
ase activity  
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
C:Comment: A large number of variants of human serum albumin have been described.  
C:Genetics:  
A:Gene: GDB:ALB  
A:Cross-references: GDB:118990; OMIM:103600  
A:Map position: 4q11-q13  
A:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridoxal  
F19-24/Domain: signal sequence #status predicted <SIG>  
F125-609/Product: serum albumin #status experimental <PRO>  
F129-202/Domain: serum albumin repeat homology <SA1>  
F166-174/Product: kinetensin #status experimental <KIP>  
F221-394/Domain: serum albumin repeat homology <SA2>  
F413-592/Domain: serum albumin repeat homology <SA3>  
F277-Binding site: copper (His) #status predicted  
F277-86-99-115-114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,  
F214/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental  
Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKQEPERNE 9  
DB 116 AKQEPERNE 124  
RESULT 4  
JC5838  
albumin - Mongolian jird  
C:Species: Meriones unguiculatus (Mongolian jird)  
C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
C:Accession: JC5838  
R:Yoshida, K.; Seto-Ohashima, A.; Sinohara, H.  
DNA Res. 4, 351-354, 1997  
A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the  
A:Reference number: JC5838; MUID:98116663; PMID:9455485  
A:Accession: JC5838  
A:Molecule type: mRNA  
A:Residues: 1-609 <YOS>  
A:Cross-references: DDBJ:AB006197; NID:92317277; PIDN:BAA21765.1; PID:92317278  
A:Experimental source: liver  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 100.0%; Score 47; DB 2; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
117 AKQEPERNE 125

Db

RESULT 5

A05139

serum albumin - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 05-Jun-1987 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000

C:Accession: A05139; 148638

R:Minghetti, P.P.; Law, S.W.; Dugaiczky, A.

Mol. Biol. Evol. 2, 347-358, 1985

A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes

A:Reference number: A93055; MUID:88216123; PMID:2452956

A:Accession: A05139

A:Molecule type: mRNA

A:Residues: 1-418 <MIN>

A:Cross-references: GB:ML6111; NID:9191764; PIDN:AAA37190.1; PID:G191765

R:Boccaccio, C.; Deschattrette, J.; Meunier-Rotival, M.

Gene 88, 181-186, 1990

A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the human genome

A:Reference number: 148638; MUID:90269606; PMID:1971802

A:Accession: 148638

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 379-453 <BOC>

A:Cross-references: EMBL:X13060; NID:952939; PIDN:CAA1459.1; PID:G899334

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>

F:123-286/Domain: serum albumin repeat homology <SA2>

F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 91.5%; Score 43; DB 2; Length 453;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
19 KOEPERNE 26

Db

RESULT 6

AB05

serum albumin precursor [validated] - bovine

N:Alternate names: 67K protein; preproalbumin

C:Species: Bos primigenius taurus (cattle)

C>Date: 24-Apr-1984 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000

C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94551

R:Holowachuk, E.W.; Stoltzberg, J.K.; Reed, R.G.; Peters Jr., T.

Submitted to the EMBL Data Library, August 1991

A:Description: Bovine serum albumin: cDNA sequence and expression.

A:Reference number: A38885

A:Accession: A38885

A:Molecule type: mRNA

A:Residues: 1-607 <HOL>

A:Cross-references: EMBL:M73215

R:Hirayana, K.; Akashi, S.; Furiya, M.; Fukuhara, K.

Biochem. Biophys. Res. Commun. 173, 639-646, 1990

A:Title: Rapid confirmation and revision of the primary structure of bovine serum albumin

A:Reference number: A36401; MUID:91083649; PMID:2260975

A:Accession: A36401

A:Molecule type: protein

A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>

R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.

Eur. J. Biochem. 98, 477-485, 1979

A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.

A:Reference number: A91258; MUID:80024278; PMID:488109

A:Accession: A91258

A:Molecule type: protein

A:Residues: 1-32 <WAG>

R:Hsieh, J.C.; Lin, F.P.; Tam, M.F.

Anal. Biochem. 170, 1-8, 1988

A:Title: Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel

A:Reference number: A60808; MUID:88267456; PMID:3389500

A:Accession: B60808

A:Molecule type: protein

A:Residues: 25-41 <HSI>

R:Strawich, E.; Glimcher, M.J.

Eur. J. Biochem. 191, 47-56, 1990

A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is albumin

A:Reference number: S10780; MUID:90336641; PMID:2379503

A:Accession: S10780

A:Molecule type: protein

A:Residues: 25-41, 'H', 43-57, 59-64 <STR>

R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.

J. Immunol. 143, 1680-1684, 1989

A:Title: Structures of histamine-releasing peptides formed by the action of acid protease

A:Reference number: A45800; MUID:89341406; PMID:2474609

A:Accession: D45800

A:Molecule type: protein

A:Residues: 163-172 <CAR>

R:Carraway, R.E.; Mitra, S.P.; Cochran, D.E.

J. Biol. Chem. 262, 5968-5973, 1987

A:Title: Structure of a biologically active neurotensin-related peptide obtained from pig

A:Reference number: A26693; MUID:87194805; PMID:2437111

A:Accession: A26693

A:Molecule type: protein

A:Residues: 165-172, 'L', 'CA2'

R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.

Biochem. J. 191, 867-868, 1980

A:Title: Sequence of residues 400-403 of bovine serum albumin.

A:Reference number: A90309; MUID:82023364; PMID:7283978

A:Accession: A90309

A:Molecule type: protein

A:Residues: 402-433 <REE>

R:Brown, J.R.

Fed. Proc. 34, 591, 1975

A:Title: Structure of bovine serum albumin.

A:Reference number: A91458

A:Accession: A91458

A:Molecule type: protein

A:Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'I'

R:Brown, J.R.

submitted to the Atlas, April 1975

A:Reference number: A94551

A:Accession: A94551

A:Molecule type: protein

A:Residues: 190-195 <BR2>

R:Brown, J.R.

Fed. Proc. 33, 1389, 1974

A:Reference number: A91457

A:Contents: annotation; disulfide bonds

R:Werlen, R.C.; Offord, R.E.; Rose, K.

Biochem. J. 302, 907-911, 1994

A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC 3.4.21.1)

A:Reference number: S55232; MUID:95031935; PMID:7945219

A:Accession: S55232

A>Status: preliminary

A:Molecule type: protein

A:Residues: 529-536; 569-572 <WER>

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; copper binding; duplication; plasma

F:1-18/Domain: signal sequence #status experimental <SIG>

F:13-24/Domain: propeptide #status experimental <PRO>

F:23-607/Product: serum albumin #status experimental <MPT>

F:23-201/Domain: serum albumin repeat homology <SA1>

F:220-393/Domain: serum albumin repeat homology <SA2>

F:412-591/Domain: serum albumin repeat homology <SA3>

F:27/Binding site: copper (His) #status predicted

F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 'I'

Query Match 91.5%; Score 43; DB 1; Length 607;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
DB 117 KOEPERNE 124

RESULT 7  
ABSHS  
serum albumin precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S06936  
R:Brown, W.M.; Dziugielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A:Reference number: S06936; MUID:90098888; PMID:2602160  
A:Accession: S06936  
A:Molecule type: mRNA  
A:Residues: 1-607 <BRO>  
A:Cross-references: EMBL:X17055; NID:g1386; PIDN:CAA34903.1; PID:g1387  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4  
F:263/Binding site: bilirubin (lys) #status predicted

Query Match 91.5%; Score 43; DB 1; Length 607;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
DB 117 KOEPERNE 124

RESULT 8  
ABHOS  
serum albumin precursor - horse  
C:Species: Equus caballus (domestic horse)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S34053  
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
Eur. J. Biochem. 215, 205-212, 1993  
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A:Reference number: S34053; MUID:93345495; PMID:8344282  
A:Accession: S34053  
A:Molecule type: mRNA  
A:Residues: 1-607 <HOA>  
A:Cross-references: GB:X74045; NID:g399671; PIDN:CAA52194.1; PID:g399672  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

F:263/Binding site: bilirubin (lys) #status predicted

Query Match 91.5%; Score 43; DB 1; Length 607;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
DB 117 KOEPERNE 124

RESULT 9  
ABPGS  
serum albumin precursor - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S01382; A61006  
R:Weinstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 9045, 1988  
A:Title: Nucleotide sequence of porcine liver albumin.  
A:Reference number: S01382; MUID:8901582; PMID:3174440  
A:Accession: S01382  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-605 <WEI>  
A:Cross-references: EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g833798  
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral  
A:Reference number: A61006; MUID:89269769; PMID:2728927  
A:Accession: A61006  
A:Molecule type: protein  
A:Residues: 23-51, 'X', 53-54; 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVW', 155 <LIM>  
A:Experimental source: dental enamel  
A:Note: albumin and other serum proteins are also found in bone  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-22/Domain: propeptide #status predicted <PRO>  
F:23-605/Product: serum albumin #status predicted <MAT>  
F:27-199/Domain: serum albumin repeat homology <SA1>  
F:218-391/Domain: serum albumin repeat homology <SA2>  
F:410-589/Domain: serum albumin repeat homology <SA3>  
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4  
F:261/Binding site: bilirubin (lys) #status predicted

Query Match 85.1%; Score 40; DB 1; Length 605;  
Best Local Similarity 87.5%; Pred. No. 6.2;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
DB 115 KOEPERNE 122

RESULT 10  
S57632  
serum albumin precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C>Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: JC4660; S57632  
R:Rilger, C.; Grigioni, F.; Hentges, F.  
Gene 169, 295-296, 1996  
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A:Reference number: JC4660; MUID:96194824; PMID:8647469  
A:Accession: JC4660  
A:Molecule type: mRNA  
A:Residues: 1-608 <HIZ>  
A:Cross-references: EMBL:X84842; NID:g986484; PIDN:CAA59279.1; PID:g986485  
A:Experimental source: liver  
C:Comment: This protein is the major protein component in plasma. It functions as a mul

```

ein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;28-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match      83.0%; Score 39; DB 2; Length 608;
Best Local Similarity 87.5%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 KQEPERNE 9
Db      117 KQEPERNE 124

RESULT 11
S59517
serum albumin precursor - monocled cobra
C;Species: Naja naja kaouthia, Naja naja siamensis (monocled cobra)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 17-Mar-2000
C;Accession: S59517
R;Wang, X.; Havsteen, B.; Hansen, H.
Biol. Chem. Hoppe-Seyler 376, 545-553, 1995
A;Title: Evidence of the coevolution of a snake toxin and its endogenous antitoxin. Clon
A;Reference number: S59517; MUID:96145734; PMID:8561913
A;Note: the source is designated as Naja naja kaouthia, but referred to as Chinese cobra
A;Accession: S59517
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-614 <WAN>
A;Cross-references: EMBL:X78598; NID:G469860; PIDN:CAA55333.1; PID:G469861
C;Superfamily: serum albumin; serum albumin repeat homology
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-614/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>

Query Match      74.5%; Score 35; DB 2; Length 614;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

QY      1 AKQEPERNE 9
Db      119 AKADPERNE 127

RESULT 12
ABCHS
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salikid, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CAS>
A;Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HAC>
A;Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039
R;Rosen, A.M.; Geller, D.M.

```

```

Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:911327
A;Accession: A13451
A;Molecule type: protein
A;Residues: 19-23; M., 23-30 <ROS>
C;Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propeptide #status predicted <PRO>
F;27-613/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397

Query Match      74.5%; Score 35; DB 1; Length 615;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

QY      1 AKQEPERNE 9
Db      119 SKADPERNE 127

RESULT 13
G96787
protein T4O12.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G96787
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96787
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-884 <STO>
A;Cross-references: GB:AE005173; NID:98778819; PIDN:AAF26771.2; GSPDB:GN00141
C;Genetics:
C;Gene: T4O12.9
A;Map position: 1
A;Map: position: 1
C;Superfamily: Schizosaccharomyces probable vacuolar protein sorting-associated protein

Query Match      74.5%; Score 35; DB 2; Length 884;
Best Local Similarity 87.5%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 KQEPERNE 9
Db      233 KQEPERNE 240

RESULT 14
A60272
IGA-specific metalloendopeptidase (EC 3.4.24.13) - Streptococcus sanguis
N;Alternate names: IGA protease; immunoglobulin A1 proteinase
C;Species: Streptococcus sanguis
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 21-Jan-2000
A;Accession: A60272
R;Gilbert, J.V.; Plaut, A.G.; Wright, A.
Infect. Immun. 59, 7-17, 1991

```

A;Title: Analysis of the immunoglobulin A protease gene of *Streptococcus sanguis*.  
A;Reference number: A60272; MUID:91100011; PMID:1987065  
A;Accession: A60272  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-1668 <GIL>  
C;Comment: This protein is shown from the start of translation of this gene as determined the start codon is shown in entry B60272.  
C;Genetics:  
A;Gene: iga  
A;Start codon: GTG  
C;Superfamily: Streptococcus sanguis Iga-specific metalloendopeptidase  
C;Keywords: hydrolase; metalloproteinase; tandem repeat; zinc

Query Match 74.5%; Score 35; DB 2; Length 1668;  
Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KOEPERN 8  
Db 353 KOEPERN 359

RESULT 15  
BOAC58  
virB10 protein - Agrobacterium tumefaciens plasmid pTiC58  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
C;Accession: S12350; S11835; S10525  
R;Shirasu, K.; Morel, P.; Kado, C.I.  
Mol. Microbiol. 4, 1153-1163, 1990  
A;Title: Characterization of the virB operon of an Agrobacterium tumefaciens Ti plasmid:  
A;Reference number: S12341; MUID:91041724; PMID:2233252  
A;Accession: S12350  
A;Molecule type: DNA  
A;Residues: 1-377 <SHI>  
A;Cross-references: EMBL:J03320; NID:gl54781; PIDN:AAA91600.1; PID:gl54792  
R;Rogowsky, P.M.; Powell, B.S.; Shirasu, K.; Lin, T.S.; Morel, P.; Zyprian, E.M.; Steck,  
Plasmid 23, 85-106, 1990  
A;Title: Molecular characterization of the vir regulon of Agrobacterium tumefaciens: com  
A;Reference number: S11825; MUID:90301800; PMID:2194232  
A;Accession: S11835  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-377 <ROG>  
A;Cross-references: EMBL:J03320; NID:gl54781; PIDN:AAA91600.1; PID:gl54792  
R;Kulda, G.A.; de Vos, G.; Owen, G.; McCaffrey, G.; Zambryski, P.  
Mol. Gen. Genet. 221, 256-266, 1990  
A;Title: The virB operon of Agrobacterium tumefaciens pTiC58 encodes 11 open reading fra  
A;Reference number: S10516; MUID:90318324; PMID:2370849  
A;Accession: S10525  
A;Molecule type: DNA  
A;Residues: 1-296, A', 298-377 <KUL>  
A;Cross-references: EMBL:X53264; NID:g39152; PIDN:CAA37363.1; PID:g39162  
C;Genetics:  
A;Gene: virB10  
A;Genome: plasmid  
C;Superfamily: tumor-inducing plasmid pTiC58 virB10 protein

Query Match 72.3%; Score 34; DB 1; Length 377;  
Best Local Similarity 66.7%; Pred. No. 57;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AKOEPERNE 9  
Db 99 ARSEPERHE 107

Search completed: April 19, 2004, 12:02:24  
Job time : 2.09695 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.65651 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 47    | 100.0       | 600    | 1     | ALBU MACMU  |
| 2          | 47    | 100.0       | 608    | 1     | ALBU RAT    |
| 3          | 47    | 100.0       | 609    | 1     | ALBU HUMAN  |
| 4          | 47    | 100.0       | 609    | 1     | ALBU MERIN  |
| 5          | 43    | 91.5        | 607    | 1     | ALBU BOVIN  |
| 6          | 43    | 91.5        | 607    | 1     | ALBU HORSE  |
| 7          | 43    | 91.5        | 607    | 1     | ALBU SHEEP  |
| 8          | 43    | 91.5        | 608    | 1     | ALBU MOUSE  |
| 9          | 40    | 85.1        | 605    | 1     | ALBU PIG    |
| 10         | 40    | 85.1        | 608    | 1     | ALBU CANFA  |
| 11         | 39    | 83.0        | 608    | 1     | ALBU FELCA  |
| 12         | 39    | 83.0        | 608    | 1     | ALBU RABIT  |
| 13         | 35    | 74.5        | 501    | 1     | A37C_DROSI  |
| 14         | 35    | 74.5        | 615    | 1     | ALBU CHICK  |
| 15         | 34    | 72.3        | 341    | 1     | SUR6_XENLA  |
| 16         | 34    | 72.3        | 377    | 1     | VIBX_AGR75  |
| 17         | 34    | 72.3        | 680    | 1     | NASP_RASIT  |
| 18         | 34    | 72.3        | 788    | 1     | NASP_HUMAN  |
| 19         | 33    | 70.2        | 251    | 1     | THIF_ECOLI  |
| 20         | 33    | 70.2        | 489    | 1     | FLIF_AQUAE  |
| 21         | 33    | 70.2        | 607    | 1     | ALB2_XENLA  |
| 22         | 33    | 70.2        | 906    | 1     | RDL4_ARATH  |
| 23         | 32    | 68.1        | 207    | 1     | VE18_NPVOP  |
| 24         | 32    | 68.1        | 215    | 1     | RS3A_THERAC |
| 25         | 32    | 68.1        | 412    | 1     | YEIT_ECC57  |
| 26         | 32    | 68.1        | 669    | 1     | STM_ENTFA   |
| 27         | 32    | 68.1        | 732    | 1     | TRIG_ECOLI  |
| 28         | 32    | 68.1        | 733    | 1     | ACE1_TRIPE  |
| 29         | 32    | 68.1        | 950    | 1     | STL2_MOUSE  |
| 30         | 32    | 68.1        | 1156   | 1     | GLH4_CABEL  |
| 31         | 32    | 68.1        | 1291   | 1     | PER1_MOUSE  |
| 32         | 32    | 68.1        | 3133   | 1     | HMCT_BOMO   |
| 33         | 32    | 68.1        | 3767   | 1     | MUA3_CABEL  |

|    |    |      |     |   |            |                    |
|----|----|------|-----|---|------------|--------------------|
| 34 | 31 | 66.0 | 143 | 1 | PSAG_HORVU | Q00327 hordeum vul |
| 35 | 31 | 66.0 | 149 | 1 | RL9_FUSNN  | Q8A110 fusobacteri |
| 36 | 31 | 66.0 | 167 | 1 | PSAG_SPIOL | P12357 spinacia ol |
| 37 | 31 | 66.0 | 193 | 1 | RHOC_HUMAN | P08134 homo sapien |
| 38 | 31 | 66.0 | 193 | 1 | RHOC_MOUSE | Q62159 mus musculu |
| 39 | 31 | 66.0 | 310 | 1 | GTRB_BPP22 | P57022 bacterioph  |
| 40 | 31 | 66.0 | 322 | 1 | RLA0_LEICH | P39096 leishmania  |
| 41 | 31 | 66.0 | 323 | 1 | RLA0_LEIIN | P39097 leishmania  |
| 42 | 31 | 66.0 | 389 | 1 | VPD_BP186  | P21679 bacterioph  |
| 43 | 31 | 66.0 | 422 | 1 | K3LI_RAT   | P43556 rattus norv |
| 44 | 31 | 66.0 | 467 | 1 | MMUP_ECOLI | Q47689 escherichia |
| 45 | 31 | 66.0 | 504 | 1 | A37C_DROME | P19487 drosophila  |

#### ALIGNMENTS

|          |  |           |      |                |
|----------|--|-----------|------|----------------|
| RESULT 1 | ALBU MACMU   | STANDARD; | PRT; | 600 AA.        |
| ID       | ALBU MACMU   | STANDARD; | PRT; | 600 AA.        |
| AC       | Q28522;  |           |      |                |
| DT       | 01-NOV-1997 (Rel. 35, Created)   |           |      |                |
| DT       | 01-NOV-1997 (Rel. 35, Last sequence update)                                |           |      |                |
| DT       | 28-FEB-2003 (Rel. 41, Last annotation update)                              |           |      |                |
| DE       | Serum albumin precursor (Fragment).  |           |      |                |
| GN       | ALB.   |           |      |                |
| OS       | Macaca mulatta (Rhesus macaque).   |           |      |                |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |           |      |                |
| OC       | Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;                  |           |      |                |
| OC       | Cercopitheidae; Macaca.  |           |      |                |
| OX       | NCBI_TaxID=9544;   |           |      |                |
| RN       | [1]  |           |      |                |
| RP       | SEQUENCE FROM N.A.   |           |      |                |
| RX       | MEDLINE=93211971; PubMed=8460152;  |           |      |                |
| RA       | Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,           |           |      |                |
| RA       | Dwilet J., Putnam F.W.   |           |      |                |
| RT       | "CDNA and protein sequence of polymorphic macaque albumins that differ     |           |      |                |
| RT       | in bilirubin binding.";  |           |      |                |
| RL       | Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).                          |           |      |                |
| CC       | !- FUNCTION: Serum albumin, the main protein of plasma, has a good         |           |      |                |
| CC       | binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,              |           |      |                |
| CC       | hormones, bilirubin and drugs. Its main function is the regulation         |           |      |                |
| CC       | of the colloidal osmotic pressure of blood.                                |           |      |                |
| CC       | !- SUBCELLULAR LOCATION: Secreted.   |           |      |                |
| CC       | !- TISSUE SPECIFICITY: Plasma.   |           |      |                |
| CC       | !- SIMILARITY: Belongs to the ALB/APP/VDB family.                          |           |      |                |
| CC       | !- SIMILARITY: Contains 3 albumin domains.                                 |           |      |                |
| CC       | This SWISS-PROT entry is copyright. It is produced through a collaboration |           |      |                |
| CC       | between the Swiss Institute of Bioinformatics and the EMBL outstation -    |           |      |                |
| CC       | the European Bioinformatics Institute. There are no restrictions on its    |           |      |                |
| CC       | use by non-profit institutions as long as its content is in no way         |           |      |                |
| CC       | modified and this statement is not removed. Usage by and for commercial    |           |      |                |
| CC       | entities requires a license agreement (See http://www.isb-sib.ch/announce/ |           |      |                |
| CC       | or send an email to license@isb-sib.ch).                                   |           |      |                |
| CC       | EMBL; M90463; AAA36906.1; -  |           |      |                |
| DR       | FIR; A47391; A47391.   |           |      |                |
| DR       | HSSP; P02768; IE7B.  |           |      |                |
| DR       | InterPro; IPR000264; Serum albumin.  |           |      |                |
| DR       | Pfam; PF00273; transport prot; 3.  |           |      |                |
| DR       | PRINTS; PD00802; SERUMALBUMIN.   |           |      |                |
| DR       | ProDom; PD002486; Serum albumin; 1.  |           |      |                |
| DR       | SMART; SM00103; ALBUMIN_3.   |           |      |                |
| DR       | PROSITE; PS00212; ALBUMIN; 3.  |           |      |                |
| KW       | Metal-binding; Lipid-binding; Repeat; Signal; Copper.                      |           |      |                |
| FT       | NON TER  | 1         |      |                |
| FT       | SIGNAL   | <1        | 10   | BY SIMILARITY. |
| FT       | PROPEP   | 11        | 16   | BY SIMILARITY. |
| FT       | CHAIN  | 17        | 600  | SERUM ALBUMIN. |
| FT       | DOMAIN   | 17        | 197  | ALBUMIN 1.     |
| FT       | DOMAIN   | 204       | 389  | ALBUMIN 2.     |
| FT       | DOMAIN   | 396       | 587  | ALBUMIN 3.     |



```

FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 295 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 47; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 108 AKQEPERNE 116

RESULT 2
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=8123722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246 (1981).
[2]
RN SEQUENCE OF 1-38, AND PROCESSING.
RP MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin
RT messenger RNA.";
RL J. Biol. Chem. 252:6846-6855 (1977).
[3]
RN SEQUENCE OF 25-222.
RP MEDLINE=78109429; PubMed=564345;
RA Isemura S., Ikenaka T.;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
RT bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48 (1978).
[4]
RN SEQUENCE OF 223-288 AND 572-608.
RP MEDLINE=76260153; PubMed=956149;
RA Isemura S., Ikenaka T.;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196 (1976).
[5]
RN SEQUENCE OF 166-174.

```

```

RC TISSUE=Plasma;
RX MEDLINE=87194805; PubMed=2437111;
RA Carraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973 (1987).
RN [6]
RP COPPER-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486 (1978).
CC -I- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -I- FUNCTION: NRP regulates fat digestion, lipid absorption, and
CC blood flow (Potential).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Plasma.
CC -I- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -I- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01222; CAA24532.1; -
DR PIR; A93872; ABRTS.
DR HSP; P02758; 1E7B.
DR InterPro; IPR000284; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM001103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608 SERUM ALBUMIN.
FT PEPTIDE 166 174 NEUROTENSIN-RELATED PEPTIDE.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT VARIANT 262 262 V -> L.
FT CONFLICT 174 174 Y -> L (IN REF. 5).
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

```

Query Match 100.0%; Score 47; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKOEPERNE 9  
Db 116 AKOEPERNE 124

RESULT 3  
ALBU HUMAN STANDARD; PRT; 609 AA.  
AC P02768; Q95574; Q13140; Q9P157; Q9P177; Q9UH53; Q9U0Z0;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RA Beattie W.G., Dugaiczky A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=8201882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczky A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the ENBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He P., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the ENBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner K.H., Shenman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madao A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573;  
RA Meloun B., Moravek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.Q.;  
RL (in) Bing D.H. (eds.);  
RT The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the ENBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=795732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Saber M.A., Stockbauer P., Moravek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,

RA Satoch C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RN Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
[18]  
RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RP MEDLINE=8345611; PubMed=2762316;  
RX Arai K., Madison J., Huss K., Ishioke N., Satoch C., Fujita M.,  
RX Neel J.V., Sakurabayashi I., Putnam F.W.;  
RA "Point substitutions in Japanese allcoalbumins";  
RT Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
[19]  
RN VARIANTS MANAU; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RP MEDLINE=90115905; PubMed=2404284;  
RX Arai K., Madison J., Shimizu A., Putnam F.W.;  
RA "Point substitutions in albumin genetic variants from Asia.";  
RT Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
[20]  
RN DESCRIPTION OF VARIANT REDHILL.  
RP MEDLINE=90115852; PubMed=2104980;  
RX Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RA "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
RT human serum albumin whose precursor has an aberrant signal peptidase  
RT cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
[21]  
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RP MEDLINE=91062352; PubMed=2247440;  
RX Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
RT Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
[22]  
RN VARIANT VENEZIA.  
RP MEDLINE=91296740; PubMed=2068071;  
RX Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
RT carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
[23]  
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RP KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
[24]  
RN VARIANT CASEBROOK ASN-519.  
RP MEDLINE=91316157; PubMed=1859851;  
RX Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
RT albumin: albumin Casebrook (494 Asp->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
[25]  
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RP MEDLINE=92190239; PubMed=1347703;  
RX Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two allcoalbumins with identical electrophoretic mobility are produced  
RT Query Match 100.0%; Score 47; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred No. 0.17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AKOEPRNE 9  
| | | | |  
Db 116 AKOEPRNE 124

REPEAT 4

```
Query Match      100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
DB      117 AKQEPERNE 125

RESULT 5
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenberg J.K., Reed R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Barry T., Power S., Gannon F.;
RC TISSUE=Liver;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RL "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE OF 1-32.
RA MEDLINE=80024278; PubMed=488109;
RA McGilivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
[6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.; of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
[7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
[8]
RP SEQUENCE OF 402-433.
RA MEDLINE=82023364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868(1980).
[9]
RP SEQUENCE OF 19-28.
RA MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine
proalbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
[10]

RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine
serum albumin by ESMS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
[11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.F.;
RT "Electroblotting onto glass-fiber filter from an analytical
isoelectrofocusing gel: a preparative method for isolating proteins
for N-terminal microsequencing.";
RL Anal. Biochem. 170:1-8(1988).
[12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
[13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RT "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389(1974).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; W73993; AAA51411.1; -
EMBL; X58989; CAA41735.1; -
EMBL; Y17769; CAA76847.1; -
EMBL; AF542068; AAN17824.1; -
HSP; P02768; 1E7B.
InterPro: IPR000264; Serum albumin.
Pfam: PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
Polymorphism. 1 18
SIGNAL 19 24
PROPEP 19 24 SERUM ALBUMIN.
FT CHAIN 25 607 ALBUMIN 1.
FT DOMAIN 25 204 ALBUMIN 2.
FT DOMAIN 211 396 ALBUMIN 3.
FT DOMAIN 403 594 COPPER (BY SIMILARITY).
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
```

FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 FT VARIANT 214 214  
 FT CONFLICT 302 302  
 FT CONFLICT 304 305  
 FT CONFLICT 324 324  
 FT CONFLICT 394 395  
 FT CONFLICT 437 437  
 FT CONFLICT 493 494  
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;

Query Match 91.5%; Score 43; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 0.97;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
 |||||  
 Db 117 KOEPERNE 124  
 |||||

RESULT 6  
 ALBU HORSE  
 ID ALBU HORSE STANDARD; PRT; 607 AA.  
 AC P35747;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serum albumin precursor (Allergen Equ c 3).  
 GN ALB.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 [1]  
 RN SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RC TISSUE=Liver;  
 RX MEDLINE=93345495; PubMed=8344282;  
 RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
 RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
 at 0.27-nm resolution.";  
 RL Eur. J. Biochem. 215:205-212(1993).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X74045; CAA52194.1; --  
 CC PIR; S34053; ABHOS.  
 CC HSSP; P02758; 1E7E.  
 CC InterPro; IPR000284; Serum\_albumin.  
 CC Pfam; PF00273; transport\_prot; 3.  
 CC PRINTS; PR00802; SERUMALBUMIN.  
 CC ProDom; PD002486; Serum\_albumin; 1.  
 CC SMART; SM00103; ALBUMIN; 3.  
 CC PROSITE; PS00212; ALBUMIN; 3.  
 CC Metal-binding; Lipid-binding; Repeat; Signal; Allergen.  
 FT SIGNAL 1 18  
 BY SIMILARITY.

FT PROPEP 19 24  
 FT CHAIN 25 607  
 FT DOMAIN 25 204  
 FT DOMAIN 211 396  
 FT DOMAIN 403 594  
 FT METAL 27 27  
 FT METAL 77 86  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 383 392  
 FT DISULFID 415 451  
 FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 SQ SEQUENCE 607 AA; 68598 MW; 258F6E830A1B90C5 CRC64;

Query Match 91.5%; Score 43; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 0.97;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9  
 |||||  
 Db 117 KOEPERNE 124  
 |||||

RESULT 7  
 ALBU SHEEP  
 ID ALBU SHEEP STANDARD; PRT; 607 AA.  
 AC P14639;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=50038888; PubMed=2602160;  
 RA Brown W.W., Dziegielewska K.M., Foreman R.C., Saunders N.R.;  
 RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";  
 RL Nucleic Acids Res. 17:10495-10495(1989).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X17055; CAA34903.1; --

DR PIR; S06936; ABSHS.  
 DR HSP; P02768; 1E7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transpore prot.; 3.  
 DR PRINTS; PRC0802; SERUMALBUMIN  
 DR PRODOM; PRC002486; Serum albumin; 1.  
 DR SMART; SMO0103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 607  
 FT DOMAIN 25 204  
 FT DOMAIN 211 396  
 FT DOMAIN 403 594  
 FT METAL 27 27  
 FT METAL 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 393 392  
 FT DISULFID 415 461  
 FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;  
 Query Match 91.5%; Score 43; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 0.97;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KOEPERNE 9  
 Db 117 KOEPERNE 124  
 RESULT 8  
 ID ALBU MOUSE STANDARD; PRT; 608 AA.  
 AC P07724; O61802;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB OR ALB1 OR ALB-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Tongue;  
 RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,  
 Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Vallaloon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 Schnur A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RN SEQUENCE OF 99-516 FROM N.A.  
 RX MEDLINE=88216123; PubMed=2452956;  
 RA Minghetti P.P., Law S.W., Dugaiczky A.;  
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that  
 of pseudogenes";  
 RL Mol. Biol. Evol. 2:347-358(1985).  
 RN [5]  
 RN SEQUENCE OF 477-551 FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=90269606; PubMed=1971802;  
 RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;  
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat  
 located in the mouse serum albumin-encoding gene";  
 RL Gene 88:181-186(1990).  
 RN [6]  
 RN SEQUENCE OF 25-44.  
 RC TISSUE=Liver;  
 RX MEDLINE=93162044; PubMed=1286668;  
 RA Giometti C.S., Taylor J., Tollaksen S.L.;  
 RT "Mouse liver protein database: a catalog of proteins detected by two-  
 dimensional gel electrophoresis";  
 RL Electrophoresis 13:970-991(1992).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way



P49822; Q77705; Q9TSZ4;  
01-OCT-1996 (Rel. 34, Created)  
16-OCT-2001 (Rel. 40, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Can f 3).  
GN ALB.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Beagle; TISSUE=Liver;  
RA Hilger C.;  
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=20148667; PubMed=10669848;  
RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,  
VA Valenta R., Spitzauer S.;  
RT "Escherichia coli expression and purification of recombinant dog  
albumin, a cross-reactive animal allergen.";  
RL J. Allergy Clin. Immunol. 105:279-285(2000).  
RN [3]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=75011422; PubMed=4414013;  
RA Dixon J.W., Sarkar B.;  
RT "Isolation, amino acid sequence and copper(II)-binding properties of  
peptide (1-24) of dog serum albumin.";  
RL J. Biol. Chem. 249:5872-5877(1974).  
RN [4]  
RP SEQUENCE OF 25-38.  
RC TISSUE=Heart;  
RX MEDLINE=98183340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2PAGE and the two-dimensional gel electrophoresis database of  
dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
RN [5]  
RP SEQUENCE OF 215-478 FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=94201492; PubMed=7512102;  
RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P.,  
Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
RT "Molecular characterization of dog albumin as a cross-reactive  
allergen.";  
RL J. Allergy Clin. Immunol. 93:614-627(1994).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -1- SUBCELLULAR LOCATION: Plasma.  
CC -1- ALLERGEN: Causes an allergic reaction in human.  
CC -1- SIMILARITY: Belongs to the ALB/APP/VBP family.  
CC -1- SIMILARITY: Contains 3 albumin domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL; A1133489; CAB64867.1; -;  
CC EMBL; Y17737; CAAY6841.1; -;  
CC EMBL; S72946; AAB30434.1; -;  
CC HSP; P02768; 1E7B.  
CC HSC-2PAGE; P49822; DOG.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport\_prot; 3.

PRINTS; P00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
GN ALB.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608  
FT DOMAIN 25 205  
FT DOMAIN 212 397  
FT DOMAIN 404 595  
FT METAL 27 27  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 289 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 340 385  
FT DISULFID 384 393  
FT DISULFID 416 462  
FT DISULFID 461 472  
FT DISULFID 485 501  
FT DISULFID 500 511  
FT DISULFID 538 583  
FT DISULFID 582 591  
FT CONFLICT 1 26  
FT CONFLICT 146 146  
FT CONFLICT 206 206  
FT CONFLICT 349 349  
FT CONFLICT 359 359  
FT CONFLICT 448 448  
FT CONFLICT 474 474  
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;  
Query Match 85.1%; Score 40; DB 1; Length 608;  
Best Local Similarity 87.5%; Pred. No. 3.7; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 0;  
QY 2 KOEPERNE 9  
DB 117 KOEPERNE 124  
RESULT 11  
ALBU\_FELCA STANDARD; PRT; 608 AA.  
ID ALBU\_FELCA  
AC P49064;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Fel d 2).  
GN ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96194824; PubMed=8647469;  
RA Hilger C., Grigioni F., Kohnen M., Kentges F.;  
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";  
RL Gene 169:295-296(1996).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Plasma.





OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7240;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=St. Lucia;  
 RX MEDLINE=9250256; PubMed=10231575;  
 RA Tataronkov A., Saez A.G., Ayala F.J.,  
 RT "A compact gene cluster in Drosophila: the unrelated Cs gene is  
 RT compressed between duplicated and and Ddc.";  
 RL Gene 231:111-120(1999).  
 CC -1- FUNCTION: Has a nonvital function (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AF091327; AAC67579.1;  
 DR FlyBase; FBgn0025651; Dsim\CG10561.  
 DR InterPro; IPR002937; Amino oxidase.  
 DR InterPro; IPR000205; NAD\_ox.  
 DR Pfam; PF01593; Amino oxidase; 1.  
 FT NON\_TER 1  
 FT 352 AEQPERN 359  
 SQ SEQUENCE 501 AA; 56018 MW; 1F140F1E45FC1249 CRC64;  
 Query Match 74.5%; Score 35; DB 1; Length 501;  
 Best Local Similarity 75.0%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKQEPERN 8  
 Db 352 AEQPERN 359  
 RESULT 14  
 ID ALBU CHICK STANDARD; PRT; 615 AA.  
 AC P19121;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).  
 GN ALB.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Cassidy A.I., Salkild C.K., Baverstock P., Wallace J.C.;  
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-28 FROM N.A.  
 RX MEDLINE=83161037; PubMed=6187737;  
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;  
 RT "The 5' noncoding and flanking regions of the avian very low density  
 RT apolipoprotein II and serum albumin genes. Homologies with the egg  
 RT white protein genes";  
 RL J. Biol. Chem. 258:4556-4564(1983).  
 RN [3]  
 RP SEQUENCE OF 19-30.  
 RX MEDLINE=78019943; PubMed=911327;  
 RA Rosen A.M., Geller D.M.;  
 RT "Chicken microsomal albumin: amino terminal sequence of chicken  
 RT proalbumin.";

RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).  
 [4]  
 RN ALLERGENIC PROPERTIES.  
 RX MEDLINE=21381307; PubMed=11488669;  
 RA Quirce S., Matanon F., Umpierrez A., de las Heras M.,  
 RA Fernandez-Caldas E., Sastre J.;  
 RT "Chicken serum albumin (Gal d 5\*) is a partially heat-labile inhalant  
 RT and food allergen implicated in the bird-egg syndrome.";  
 RL Allergy 56:754-762(2001).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- ALLERGEN: Causes an allergic reaction in human. Binds IgE.  
 CC Partially heat-labile allergen that may cause both respiratory and  
 CC food-allergy symptoms in patients with the bird-egg syndrome.  
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X60688; CAA43098.1;  
 DR EMBL; V00381; CAA23680.1;  
 DR PIR; S15571; ABCS.  
 DR HSP; P02758; 1E7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 23  
 FT CHAIN 24 615  
 FT DOMAIN 24 209  
 FT DOMAIN 216 401  
 FT DOMAIN 408 559  
 FT METAL 30 30  
 FT DISULFID 80 89  
 FT DISULFID 102 118  
 FT DISULFID 117 128  
 FT DISULFID 152 197  
 FT DISULFID 196 205  
 FT DISULFID 228 274  
 FT DISULFID 273 281  
 FT DISULFID 293 307  
 FT DISULFID 306 317  
 FT DISULFID 344 359  
 FT DISULFID 368 397  
 FT DISULFID 420 466  
 FT DISULFID 465 476  
 FT DISULFID 489 505  
 FT DISULFID 504 515  
 FT DISULFID 542 587  
 FT DISULFID 586 595  
 FT CARBOHYD 500 500  
 FT CONFLICT 24 24  
 FT SEQUENCE 615 AA; 69918 MW; E59E4BBAC066C6 CRC64;  
 Query Match 74.5%; Score 35; DB 1; Length 615;  
 Best Local Similarity 66.7%; Pred. No. 34;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKQEPERN 9

Db 119 SKADPERNE 127

## RESULT 15

SUR6 XENLA  
ID SUR6 XENLA STANDARD; PRT; 341 AA.  
AC QSI80;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 18-OCT-2001 (Rel. 40, Last sequence update)  
DT 18-OCT-2001 (Rel. 40, Last annotation update)  
DE Surfeit locus protein 6 homolog.  
GN SURF6.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xeropodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Wolff C.M.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Involved in a nucleolar function (By similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).

CC -!- SIMILARITY: Belongs to the SURF6 family.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; AJ276843; CAB98154.1; -.

DR InterPro; IPR007019; SURF6.

DR Pfam; PF04935; SURF6; 1.

KW Nuclear protein.

SQ SEQUENCE 341 AA; 39901 MW; 2E0A06920D653A0 CRC64;

## Query Match

Best Local Similarity 72.3%; Score 34; DB 1; Length 341;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

Db 152 AAQEPPEGNE 160

Search completed: April 19, 2004, 11:52:49

Job time : 1.65651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_todent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 47    | 100.0       | 396    | Q81UK7    | Q81UK7 homo sapien |
| 2          | 43    | 91.5        | 205    | 11 Q8CG74 | Q8CG74 mus musculu |
| 3          | 43    | 91.5        | 576    | 11 Q8C7C7 | Q8C7C7 mus musculu |
| 4          | 43    | 91.5        | 608    | 5 Q95VB7  | Q95VB7 schistosoma |
| 5          | 43    | 91.5        | 608    | 11 Q8C7H3 | Q8C7H3 mus musculu |
| 6          | 39    | 83.0        | 584    | 6 Q7YSG3  | Q7YSG3 felis silve |
| 7          | 36    | 76.6        | 115    | 5 Q8MR24  | Q8MR24 drosophila  |
| 8          | 36    | 76.6        | 203    | 5 Q8IN20  | Q8IN20 drosophila  |
| 9          | 36    | 76.6        | 439    | 5 Q917X6  | Q917X6 drosophila  |
| 10         | 36    | 76.6        | 456    | 5 Q46078  | Q46078 drosophila  |
| 11         | 35    | 74.5        | 414    | 2 Q7WSR5  | Q7WSR5 arthrobacte |
| 12         | 35    | 74.5        | 443    | 11 Q8BI09 | Q8BI09 mus musculu |
| 13         | 35    | 74.5        | 566    | 11 Q81WGS | Q81WGS mus musculu |
| 14         | 35    | 74.5        | 614    | 13 Q81134 | Q81134 naja naja   |
| 15         | 35    | 74.5        | 884    | 10 Q9LQS9 | Q9LQS9 arabidopsis |
| 16         | 35    | 74.5        | 1853   | 2 Q33761  | Q33761 streptococc |

|    |    |      |      |    |        |                    |
|----|----|------|------|----|--------|--------------------|
| 17 | 35 | 74.5 | 1854 | 2  | Q33764 | Q33764 streptococc |
| 18 | 35 | 74.5 | 1854 | 2  | Q33759 | Q33759 streptococc |
| 19 | 35 | 74.5 | 1874 | 2  | Q33765 | Q33765 streptococc |
| 20 | 35 | 74.5 | 1874 | 2  | Q33763 | Q33763 streptococc |
| 21 | 35 | 74.5 | 1878 | 2  | Q59986 | Q59986 streptococc |
| 22 | 35 | 74.5 | 1914 | 2  | Q33762 | Q33762 streptococc |
| 23 | 34 | 72.3 | 318  | 4  | Q9P1N1 | Q9P1N1 homo sapien |
| 24 | 34 | 72.3 | 375  | 2  | Q66290 | Q66290 agrobacteri |
| 25 | 34 | 72.3 | 421  | 11 | Q88993 | Q88993 mus musculu |
| 26 | 34 | 72.3 | 448  | 11 | Q99K39 | Q99K39 mus musculu |
| 27 | 34 | 72.3 | 684  | 4  | Q9Y6X4 | Q9Y6X4 homo sapien |
| 28 | 34 | 72.3 | 750  | 11 | Q7TM96 | Q7TM96 rattus norv |
| 29 | 34 | 72.3 | 772  | 11 | Q99MD9 | Q99MD9 mus musculu |
| 30 | 34 | 72.3 | 773  | 11 | Q35499 | Q35499 mus musculu |
| 31 | 34 | 72.3 | 2011 | 4  | Q920A9 | Q920A9 homo sapien |
| 32 | 33 | 70.2 | 65   | 4  | Q9BUN3 | Q9BUN3 homo sapien |
| 33 | 33 | 70.2 | 77   | 15 | Q7ZPH8 | Q7ZPH8 human immun |
| 34 | 33 | 70.2 | 178  | 12 | Q7TD10 | Q7TD10 cestrum yel |
| 35 | 33 | 70.2 | 245  | 16 | Q8X621 | Q8X621 escherichia |
| 36 | 33 | 70.2 | 256  | 16 | Q8FB79 | Q8FB79 escherichia |
| 37 | 33 | 70.2 | 284  | 16 | Q7VUL4 | Q7VUL4 bordetella  |
| 38 | 33 | 70.2 | 343  | 10 | Q8S1Z8 | Q8S1Z8 arabidopsis |
| 39 | 33 | 70.2 | 399  | 16 | Q8K6Q6 | Q8K6Q6 bacillus ha |
| 40 | 33 | 70.2 | 407  | 16 | Q91323 | Q91323 pseudomonas |
| 41 | 33 | 70.2 | 433  | 16 | Q9HTB4 | Q9HTB4 pseudomonas |
| 42 | 33 | 70.2 | 503  | 5  | Q25777 | Q25777 plasmodium  |
| 43 | 33 | 70.2 | 521  | 5  | Q9VDD3 | Q9VDD3 drosophila  |
| 44 | 33 | 70.2 | 555  | 4  | Q8NA99 | Q8NA99 homo sapien |
| 45 | 33 | 70.2 | 713  | 16 | Q8UJI6 | Q8UJI6 agrobacteri |

#### ALIGNMENTS

RESULT 1  
Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7;  
DT 01-MAR-2003 (TREMREL. 23, Created)  
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)  
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
DE Similar to serum albumin precursor.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035969; AAH35969.1; -  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 2.  
DR PRINTS; PRO0802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 47; DB 4; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.43; Indels 0; Gaps 0;  
Matches 9; Mismatches 0;

Qy 1 AKQEPERNE 9  
Db 116 AKQEPERNE 124

RESULT 2  
Q8CG74

ID Q8CG74 PRELIMINARY; PRT; 205 AA.  
AC Q8CG74;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Albumin (fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvevTACfBr;  
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ277794; CAC81903.1;  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 1.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 1.  
FT NON TER 205  
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;  
Query Match 91.5%; Score 43; DB 11; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KQEPERNE 9  
DB 117 KQEPERNE 124  
RESULT 3  
Q8C7C7 PRELIMINARY; PRT; 576 AA.  
ID Q8C7C7;  
AC Q8C7C7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Albumin 1 (Fragment).  
GN ALB1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Thymus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK050644; BAC34360.1; -.  
DR MGD; MGI:87991; Alb1.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
FT NON TER 1  
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;  
Query Match 91.5%; Score 43; DB 11; Length 576;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KQEPERNE 9  
DB 117 KQEPERNE 124  
RESULT 4  
Q95VB7 PRELIMINARY; PRT; 608 AA.  
ID Q95VB7;  
AC Q95VB7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Albumin.  
OS Schistosoma mansoni (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
OC Schistosomatoidea; Schistosomatidae; Schistosoma.  
OX NCBI\_TaxID=6183;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;  
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg  
RT component in murine infection with Schistosoma mansoni."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF418550; AAL08579.1; -.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
DR SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;  
Query Match 91.5%; Score 43; DB 5; Length 608;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKQEPERNE 9  
DB 116 AKQEPERNE 124  
RESULT 5  
Q8C7H3 PRELIMINARY; PRT; 608 AA.  
ID Q8C7H3;  
AC Q8C7H3;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Albumin 1.  
GN ALB1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK050248; BAC34145.1; -.  
DR MGD; MGI:87991; Alb1.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.

DR InterPro: IPR000264; Serum\_albumin.  
DR Pfam: PF00273; transprot\_prot; 3.  
DR PRINTS: PR00802; SERUMALBUMIN.  
DR PRODOM: PD002486; Serum\_albumin; 1.  
DR SMART: SM00103; ALBUMIN; 3.  
DR PROSITE: PS00212; ALBUMIN; 3.  
SQ SEQUENCE 608 AA; 68722 MW; 292F600ED3A61B4 CRC64;  
  
Query Match 91.5%; Score 43; DB 11; Length 608;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KQEPERNE 9  
|:|||||  
Db 117 KQEPERNE 124  
  
RESULT 6  
QYXSG3 PRELIMINARY; PRT; 584 AA.  
AC QYXSG3  
DT 01-OCT-2003 (TReMBLrel. 25, Created)  
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Albumin (fragment).  
GN ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,  
RA Rumpold H., Valenta R., Spitzauer S.;  
RT "Escherichia coli expression and purification of recombinant cat  
albumin: IgE recognition, induction of basophil activation and  
lymphoproliferative responses in atopic patients.";  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ487677; CAD32275.1; -  
FT NON TER 1  
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;  
  
Query Match 83.0%; Score 39; DB 6; Length 584;  
Best Local Similarity 87.5%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KQEPERNE 9  
|:|||||  
Db 93 KQEPERNE 100  
  
RESULT 7  
QYXSG3 PRELIMINARY; PRT; 115 AA.  
AC QYXSG3  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)  
DE SD11171p.  
GN BCDNA:SD11171.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.N.,

RA Celniker S.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY119187; AAMS1047.1; -  
DR FlyBase: FBGN0062887; BCDNA:SD11171.  
SQ SEQUENCE 115 AA; 13534 MW; F680D64C7C967AEC CRC64;  
  
Query Match 76.6%; Score 36; DB 5; Length 115;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KQEPERNE 9  
|:|||||  
Db 34 KQEPERNE 41  
  
RESULT 8  
QYXSG3 PRELIMINARY; PRT; 203 AA.  
AC QYXSG3  
DT 01-MAR-2003 (TReMBLrel. 23, Created)  
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE CG31174-PA.  
GN CG31174.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Frankoch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K.J., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skuiski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swires R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.H.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam T.C., Kruse D., Li P., Matel B., Moshrefi A., Nunoo J.,  
RA Pacleeb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RA "Sequencing of Drosophila melanogaster genome";  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celniker S.E.,  
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RA "Annotation of Drosophila melanogaster genome";  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE003736; AAN13887.1; -;  
DR FlyBase; FBgn0051174; CG31174.  
SQ SEQUENCE 203 AA; 23859 MW; 3C3BEF93F3434B93 CRC64;  
  
Query Match 76.6%; Score 36; DB 5; Length 203;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KOEPERNE 9  
:|||||:  
DB 34 KREPERNQ 41  
  
RESULT 9  
Q917X6 PRELIMINARY; PRT; 439 AA.  
ID Q917X6  
AC Q917X6; (TREMELrel. 16, Created)  
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMELrel. 19, Last annotation update)  
DE EG:39E1.1 protein (LD42227p).  
GN EG:39E1.1 OR CG11596.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman I.J., Hernandez J.P., Houck J.,  
RA Hostin D., Hostin K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mersiklova G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzay D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleeb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE003422; AAG22387.1; -;  
DR EMBL: AY051951; AAK93375.1; -;  
DR FlyBase; FBgn0023522; EG:39E1.1.  
SQ SEQUENCE 439 AA; 50504 MW; 883DE54CC8D8B5F7 CRC64;  
  
Query Match 76.6%; Score 36; DB 5; Length 439;  
Best Local Similarity 75.0%; Pred. No. 74;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KOEPERNE 9  
:|||||:  
DB 408 EQOEPERNE 415  
  
RESULT 10  
O46078 PRELIMINARY; PRT; 456 AA.  
ID O46078  
AC O46078; Q9W540;  
DT 01-JUN-1998 (TREMELrel. 06, Created)  
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE EG:39E1.1 protein.  
GN EG:39E1.1 OR CG11596.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

|  |  |
|--|--|
| Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,          |  |
| Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,   |  |
| Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  |  |
| Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,   |  |
| Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,    |  |
| Beeson K.V., Benos P.V., Birman B.P., Bhandari D., Bolshakov S.,       |  |
| Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,         |  |
| Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  |  |
| Cherry J.M., Cavley S., Dahke C., Davenport L.B., Davies P.,           |  |
| de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,      |  |
| Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., |  |
| Dubin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  |  |
| Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,       |  |
| Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,           |  |
| Harris N.L., Harvey D., Heiman T.H., Hernandez J.R., Houck T.,         |  |
| Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,          |  |
| Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., |  |
| Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,       |  |
| Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,       |  |
| Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,           |  |
| Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,        |  |
| Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,     |  |
| Nelson D.R., Nelson K.A., Nixon K., Nusserd J.R., Pacleb J.M.,         |  |
| Palaizolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,   |  |
| Reibert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,       |  |
| Shue B.C., Siden-Kiamos I., Simpson M., Skubecki M.P., Smith T.,       |  |
| Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,             |  |
| Svirskas R., Tecor C., Turner K., Venter E., Wang A.H., Wang X.,       |  |
| Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,            |  |
| Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,      |  |
| Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,   |  |
| Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., |  |
| Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.                        |  |
| "The genome sequence of <i>Drosophila melanogaster</i> ";              |  |
| Science 287:2185-2195 (2000).  |  |
| [2]  |  |
| SEQUENCE FROM N.A.   |  |
| STRAIN=OREGON-R;   |  |
| MEDLINE=20196011; PubMed=10731137;                                     |  |
| Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D., Barrell B., |  |
| Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E., Drano S.,       |  |
| Gloux S., Lelaure V., Mottier S., Gallibert F., Borkova D., Minana B., |  |
| Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,                 |  |
| Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,         |  |
| Modellell J., Peter A., Schottler P., Werner M., Mourikoti F.,         |  |
| Beirnt N., Dowe G., Schafer U., Jackle H., Bucheton A.,                |  |
| Callister D.M., Campbell L.A., Darlamitsov A., Henderson N.S.,         |  |
| McMillan P.J., Sales C., Tait E.A., Valenti P., Saunderson R.D.,       |  |
| Glover D.M.;   |  |
| "From sequence to chromosome: the tip of the X chromosome of <i>D.</i> |  |
| <i>melanogaster</i> .";  |  |
| Science 287:2220-2222 (2000).  |  |
| -!- ALTERNATIVE PRODUCTS:  |  |
| Name=Long;   |  |
| IsId=O4D6078-1; Sequence=Displayed;                                    |  |
| Name=Short;  |  |
| IsId=O4D6078-2; Sequence=VSP_050174;                                   |  |
| EMBL; AF003422; AAFA5697.1; -  |  |
| EMBL; AF003422; AAFA5698.1; -  |  |
| EMBL; AL008191; CAAL5684.1; -  |  |
| FlyBase; Fggn0023522; EG:39E1.1.                                       |  |
| ALTERNATIVE splicing; Hypothetical protein.                            |  |
| VARSPLIC 1 55 Missing (in isoform Short).                              |  |
| /FTID=VSP_050174.  |  |
| SEQUENCE 456 AA; 52441 MW; CC6BCF2FBADD5954 CRC64;                     |  |
| Query Match 76.6%; Score 36; DB 5; Length 456;                         |  |
| Best Local Similarity 75.0%; Pred.No. 77;                              |  |
| Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;             |  |
| QY 2 KORPERNE 9  |  |
| . ..   |  |



## RESULT 13

Q91WGS  
ID Q91WGS PRELIMINARY; PRT; 566 AA.  
AC Q91WGS  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN 2410051C13RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC015283; AAH15283.1; -  
DR MGD; MGI:1920950; 2410051C13RIK.  
DR InterPro; IPR000644; CBS\_domain.  
DR Pfam; PF00571; CBS; 4.  
DR SMART; SM00116; CBS; 4.  
KW Hypothetical protein.  
SQ SEQUENCE 566 AA; 62997 MW; 754DE4D696102C7D CRC64;

Query Match 74.5%; Score 35; DB 11; Length 566;  
Best Local Similarity 87.5%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KOEPERNE 9

Db 181 KOEPERPE 188

## RESULT 14

Q91134  
ID Q91134 PRELIMINARY; PRT; 614 AA.  
AC Q91134;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Cobra serum albumin.  
OS Naja naja (Indian cobra).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Elapidae; Elapinae; Naja.  
OX NCBI\_TaxID=35670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=96145734; PubMed=8561913;  
RA Wang X., Hansen H., Havsteen B.;  
RT "Evidence of the coevolution of snake toxin and its enogenous  
RT antitoxin. Cloning, sequence and expression of a serum albumin cDNA of  
RT the chinese cobra";  
RL Biol. Chem. Hoppe-Seyler 376:545-553 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93343893; PubMed=8343135;  
RA Shao J., Shen H., Havsteen B.;  
RT "Purification, characterization and binding interactions of the  
RT Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP.";  
RL Biochem. J. 293:559-566 (1993).  
DR EMBL; X78598; CAA55333.1; -  
DR HSSP; P02768; 1E7B.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0004810; P:transport; IEA.  
DR InterPro; IPR000264; Serum\_albumin.

DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 614 AA; 69798 MW; 3DB2D3CC4BD8CBFD CRC64;

Query Match 74.5%; Score 35; DB 13; Length 614;  
Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

Db 119 AKADPEPNE 127

## RESULT 15

Q9LQ99  
ID Q9LQ99 PRELIMINARY; PRT; 884 AA.  
AC Q9LQ99;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE T4012.9.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chao Q., Shinn P., Brooks S., Buehler E., Dunn P., Khan S., Kim C.,  
RA Walker M., Brooks S., Altafi H., Araujo R., Conn L., Conway A.B.,  
RA Gonzalez A., Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J.,  
RA Liu S., Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V.,  
RA Yu G., Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;  
RT "Genomic sequence for Arabidopsis thaliana BAC T4012 from chromosome  
RT I.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
RA Shinn P., Altafi H., Bei B., Chin C., Chlou J., Choi E., Conn L.,  
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,  
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,  
RA Ecker J.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chlou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007396; AAF26771.2; -  
DR PIR; G96787; G96787.  
DR InterPro; IPR005378; Vps35.  
DR Pfam; PF03635; Vps35; 1.  
SQ SEQUENCE 884 AA; 100411 MW; B4BFD781BDALD048 CRC64;

Query Match 74.5%; Score 35; DB 10; Length 884;  
 Best Local Similarity 87.5%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QV 2 KQEPERNE 9  
 |||||  
 Db 233 KQEKERNE 240

Search completed: April 19, 2004, 12:00:03  
 Job time : 5.47368 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:23 ; Search time 5.29363 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47  
Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980as:\*
- 2: Geneseqp1990as:\*
- 3: Geneseqp2000as:\*
- 4: Geneseqp2001as:\*
- 5: Geneseqp2002as:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 47    | 100.0       | 116    | AAU29580 | AAU29580 Novel hum |
| 2          | 47    | 100.0       | 116    | AAU29574 | AAU29574 Novel hum |
| 3          | 47    | 100.0       | 137    | AAO12088 | AAO12088 Human pol |
| 4          | 47    | 100.0       | 143    | ABP05068 | ABP05068 Human ORF |
| 5          | 47    | 100.0       | 156    | AAU33073 | AAU33073 Novel hum |
| 6          | 47    | 100.0       | 156    | AAU33271 | AAU33271 Novel hum |
| 7          | 47    | 100.0       | 192    | AAU29875 | AAU29875 Novel hum |
| 8          | 47    | 100.0       | 195    | AAO17048 | AAO17048 Human ser |
| 9          | 47    | 100.0       | 195    | ABU10022 | ABU10022 Human ser |
| 10         | 47    | 100.0       | 204    | AAU83947 | AAU83947 Yeast cod |
| 11         | 47    | 100.0       | 236    | AAO17051 | AAO17051 Human alb |
| 12         | 47    | 100.0       | 241    | AAO16984 | AAO16984 Alpha-MSH |
| 13         | 47    | 100.0       | 241    | ABU10025 | ABU10025 Alpha-MSH |
| 14         | 47    | 100.0       | 242    | AAO16985 | AAO16985 Alpha-MSH |
| 15         | 47    | 100.0       | 244    | AAO16986 | AAO16986 Alpha-MSH |
| 16         | 47    | 100.0       | 245    | AAO16987 | AAO16987 Alpha-MSH |
| 17         | 47    | 100.0       | 245    | AAO16988 | AAO16988 Alpha-MSH |
| 18         | 47    | 100.0       | 268    | AAO16989 | AAO16989 Alpha-MSH |
| 19         | 47    | 100.0       | 268    | ABU10026 | ABU10026 Alpha-MSH |
| 20         | 47    | 100.0       | 303    | AAU14178 | AAU14178 Human ser |
| 21         | 47    | 100.0       | 327    | AAU32564 | AAU32564 Novel hum |
| 22         | 47    | 100.0       | 327    | AAU29942 | AAU29942 Novel hum |
| 23         | 47    | 100.0       | 327    | AAU32995 | AAU32995 Novel hum |
| 24         | 47    | 100.0       | 327    | AAU33287 | AAU33287 Novel hum |
| 25         | 47    | 100.0       | 373    | AAU90387 | AAU90387 N-termina |

|    |    |       |     |   |          |                    |
|----|----|-------|-----|---|----------|--------------------|
| 26 | 47 | 100.0 | 388 | 1 | AAU90389 | AAU90389 N-termina |
| 27 | 47 | 100.0 | 389 | 1 | AAU90390 | AAU90390 N-termina |
| 28 | 47 | 100.0 | 390 | 1 | AAU90391 | AAU90391 N-termina |
| 29 | 47 | 100.0 | 407 | 1 | AAU90392 | AAU90392 N-termina |
| 30 | 47 | 100.0 | 500 | 7 | ADD32019 | ADD32019 Heterolog |
| 31 | 47 | 100.0 | 584 | 6 | ABG72381 | ABG72381 Mature hu |
| 32 | 47 | 100.0 | 585 | 1 | AAU90344 | AAU90344 Sequence  |
| 33 | 47 | 100.0 | 585 | 1 | AAU90388 | AAU90388 Mature hu |
| 34 | 47 | 100.0 | 585 | 1 | AAU91422 | AAU91422 Human nor |
| 35 | 47 | 100.0 | 585 | 1 | AAU05318 | AAU05318 Human ser |
| 36 | 47 | 100.0 | 585 | 2 | AAU08457 | AAU08457 Human ser |
| 37 | 47 | 100.0 | 585 | 2 | AAU26207 | AAU26207 Human ser |
| 38 | 47 | 100.0 | 585 | 2 | AAU26362 | AAU26362 Synthetic |
| 39 | 47 | 100.0 | 585 | 2 | AAU20029 | AAU20029 Human ser |
| 40 | 47 | 100.0 | 585 | 2 | AAU80301 | AAU80301 Human ser |
| 41 | 47 | 100.0 | 585 | 2 | AAU20111 | AAU20111 HSA prote |
| 42 | 47 | 100.0 | 585 | 2 | AAU59841 | AAU59841 Mature pr |
| 43 | 47 | 100.0 | 585 | 3 | AAU84873 | AAU84873 Amino aci |
| 44 | 47 | 100.0 | 585 | 3 | AAU83946 | AAU83946 Yeast cod |
| 45 | 47 | 100.0 | 585 | 4 | AAU52567 | AAU52567 Mature hu |

## ALIGNMENTS

## RESULT 1

AAU29580  
ID AAU29580 standard; protein; 116 AA.

XX AC AAU29580;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #71.

XX KW Human; vaccination; gene therapy; nutritional supplement;

XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX FN WC200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy.

XX PS Claim 20; Page 168; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 116 AA;

Query Match 100.0%; Score 47; DB 4; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;  
 QY 1 AKQEPERNE 9  
 |||||  
 Db 6 AKQEPERNE 14

RESULT 2  
 AAU29574  
 ID AAU29574 standard; protein; 116 AA.  
 AC AAU29574;  
 XX  
 DT 18-DEC-2001 (first entry)  
 DE Novel human secreted protein #65.

XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.

XX WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008656.  
 XX 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.

XX Claim 20; Page 167; 765pp; English.  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 116 AA;

Query Match 100.0%; Score 47; DB 4; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;  
 QY 1 AKQEPERNE 9  
 |||||  
 Db 6 AKQEPERNE 14

RESULT 3  
 AAO12088  
 ID AAO12088 standard; protein; 137 AA.  
 AC AAO12088;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polypeptide SEQ ID NO 25980.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.  
 OS WO200164835-A2.  
 XX 07-SEP-2001.  
 XX 26-FEB-2001; 2001WO-US004927.  
 XX 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-514838/56.  
 DR N-PSDB; AAI92019.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 25980; 1399pp + Sequence Listing; English.  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 137 AA;

Query Match 100.0%; Score 47; DB 4; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
 |||||  
 Db 121 AKQEPERNE 129

RESULT 4

ABP05068  
ID ABP05068 standard; protein; 143 AA.  
XX  
AC ABP05068;  
XX  
DI 24-JUN-2002 (first entry)  
XX  
DE Human ORFX protein sequence SEQ ID NO:10118.  
XX  
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.  
XX  
OS Homo sapiens.  
XX  
FN WO200192523-A2.  
XX  
FD 06-DEC-2001.  
XX  
PF 29-MAY-2001; 2001WO-US010836.  
XX  
PR 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shinkets RA, Leach MD;  
XX  
DR WPI; 2002-106308/14.  
DR N-PSDB; ABN20820.  
XX  
PT Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX  
PS Disclosure; SEQ ID NO 10118; 1037pp; English.  
XX  
CC The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification)). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 143 AA;

Query Match 100.0%; Score 47; DB 5; Length 143;  
Best Local Similarity 100.0%; Pred. No. 0.61;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
DB 111 AKQEPERNE 119  
XX  
RESULT 5  
AAU33073  
ID AAU33073 standard; protein; 156 AA.  
XX  
AC AAU33073;  
XX  
DI 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3564.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
FN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR 26-JAN-2001; 2001US-00770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
PS Claim 20; Page 704; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 156 AA;

Query Match 100.0%; Score 47; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
DB 47 AKQEPERNE 55  
XX  
RESULT 6  
AAU33271  
ID AAU33271 standard; protein; 156 AA.

XX AC AAU3327L;  
 XX DT 18-DEC-2001 (first entry)  
 XX DE Novel human secreted protein #3762.  
 XX XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX OS Homo sapiens.  
 XX PN WO200179449-A2.  
 XX PD 25-OCT-2001.  
 XX PF 16-APR-2001; 2001WO-US008656.  
 XX PR 18-APR-2000; 2000US-00552929.  
 XX PR 26-JAN-2001; 2001US-00770160.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Drmanac RT;  
 XX DR WPI; 2001-611725/70.  
 XX XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX PS Claim 20; Page 751; 765pp; English.  
 XX CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX SQ Sequence 156 AA;  
 Query Match 100.0%; Score 47; DB 4; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 0.66;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKQEPERNE 9  
 |||||  
 Db 47 AKQEPERNE 55  
 RESULT 7  
 AAU29875  
 ID AAU29875 standard; protein; 192 AA.  
 XX AC AAU29875;  
 XX XX 18-DEC-2001 (first entry)  
 DT Novel human secreted protein #366.  
 XX DE Human; vaccination; gene therapy; nutritional supplement;  
 KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 OS Homo sapiens.  
 XX PN WO200179449-A2.  
 XX PD 25-OCT-2001.  
 XX PF 16-APR-2001; 2001WO-US008656.  
 XX PR 18-APR-2000; 2000US-00552929.  
 XX PR 26-JAN-2001; 2001US-00770160.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Drmanac RT;  
 XX DR WPI; 2001-611725/70.  
 XX XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX PS Claim 20; Page 206; 765pp; English.  
 XX CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX SQ Sequence 192 AA;  
 Query Match 100.0%; Score 47; DB 4; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 0.81;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKQEPERNE 9  
 |||||  
 Db 98 AKQEPERNE 106  
 RESULT 8  
 AAU17048  
 ID AAU17048 standard; protein; 195 AA.  
 XX AC AAU17048;  
 XX AC AAU17048;  
 DT 29-MAY-2002 (first entry)  
 XX DE Human serum albumin (1-195) SEQ ID NO: 56.  
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
 KW antidiabetic; antibacterial; dermatological; antiporiatic;  
 KW diabetes; uveitis; coeliac disease.  
 XX OS Homo sapiens.

XX WO200206316-A2.  
 XX 24-JAN-2002.  
 XX 16-JUL-2001; 2001WO-US022263.  
 XX 14-JUL-2000; 2000US-0218381P.  
 XX 18-AUG-2000; 2000US-0226382P.  
 XX 06-OCT-2000; 2000US-0238380P.  
 XX 29-DEC-2000; 2000US-0258764P.  
 XX 14-JUN-2001; 2001US-0298317P.  
 XX (ZYCO-) ZYCO INC.  
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 XX WPI; 2002-195801/25.  
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 XX stimulating hormone concatamer or its analog, for treating inflammatory  
 XX or autoimmune disorders.  
 XX Example 2; Page 46; 89pp; English.  
 XX The present invention relates to a nucleic acid comprising a sequence  
 XX encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 XX hormone (MSH) concatamer. The sequences are useful for treating an  
 XX individual suffering from, or at risk of, a disorder of the immune system  
 XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 XX present sequence is a protein described in the exemplification of the  
 XX invention  
 XX SQ Sequence 195 AA;  
 Query Match 100.0%; Score 47; DB 5; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKQEPERNE 9  
 DB 92 AKQEPERNE 100  
 RESULT 9  
 ABU10022  
 ID ABU10022 standard; protein; 195 AA.  
 XX AC ABU10022;  
 XX 31-JUL-2003 (first entry)  
 XX Human serum albumin residues 1-195.  
 XX Bladder disorder; cytostatic; antiinflammatory; immune response;  
 XX un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
 XX bladder cancer; tumour; interstitial cystitis; inflammation;  
 XX alpha-MSH concatamer; melanocyte stimulating hormone; human;  
 XX serum albumin.  
 XX Homo sapiens.  
 XX US2002193332-A1.  
 XX 19-DEC-2002.  
 XX 12-FEB-2002; 2002US-00074956.  
 XX 12-FEB-2001; 2001US-0268175P.  
 XX

PA (HEDL/) HEDLEY M L.  
 XX Hedley ML;  
 XX WPI; 2003-447327/42.  
 XX Modulating immune responses in a mammal with a bladder disorder e.g.  
 XX bladder cancer, by administering nucleic acids comprising un-methylated  
 XX CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
 XX the mammal.  
 XX Example 2; Page 9; 17pp; English.  
 XX The invention describes a method of modulating an immune response in a  
 XX mammal, comprising identifying a mammal that has or is at risk for having  
 XX a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
 XX comprising an un-methylated CpG sequence to the mammal, (b) an isolated  
 XX nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal,  
 XX or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
 XX method is useful for modulating immune response in a mammal having a  
 XX bladder disorder, where administration of (N1) results in an amelioration  
 XX of one or more symptoms of the disorder. Preferably, the method is useful  
 XX for modulating immune response in a mammal having bladder cancer (where  
 XX administration of (N1) results in a decrease in tumour size or activity),  
 XX or for modulating immune response in a mammal having interstitial  
 XX cystitis (where administration of (N1) results in a modulation of the  
 XX immune response from Th2 response to a Th1 response). The method is also  
 XX useful for modulating immune response in a mammal having bladder disorder  
 XX that is characterised by inflammation which is associated with symptoms  
 XX of interstitial cystitis or associated with a disruption of the integrity  
 XX of the bladder lining. This is the amino acid sequence of human serum  
 XX albumin residues 1-195 that can be used in the creation of melanocyte  
 XX stimulating hormone (alpha-MSH) concatamers resulting in secretion of the  
 XX fusion protein when expressed in mammalian cells  
 XX SQ Sequence 195 AA;  
 Query Match 100.0%; Score 47; DB 7; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKQEPERNE 9  
 DB 92 AKQEPERNE 100  
 RESULT 10  
 AAY83947  
 ID AAY83947 standard; protein; 204 AA.  
 XX AC AAY83947;  
 XX 28-JUL-2000 (first entry)  
 XX Yeast codon-biased recombinant HSA protein fragment HSA-I.  
 XX Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
 XX overlapping oligonucleotide; expression vector.  
 XX Homo sapiens.  
 XX Synthetic.  
 XX CN1239103-A.  
 XX 22-DEC-1999.  
 XX 17-JUN-1998; 98CN-00102506.  
 XX 17-JUN-1998; 98CN-00102506.  
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
 XX Li S, Lu D;  
 XX





CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
XX  
SQ Sequence 241 AA;  
Query Match 100.0%; Score 47; DB 5; Length 241;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKQEPERNE 9  
Db 116 AKQEPERNE 124  
|||||  
RESULT 13  
ABU10025  
ID ABU10025 standard; protein; 241 AA.  
XX  
AC ABU10025;  
XX  
DT 31-JUL-2003 (first entry)  
XX  
DE Alpha-MSH/Serum albumin fusion protein H9.  
XX  
KW Bladder disorder; cytostatic; antinflammatory; immune response;  
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
KW bladder cancer; tumour; interstitial cystitis; inflammation;  
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;  
KW serum albumin; fusion protein; H9.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN US2002193332-A1.  
XX  
PD 19-DEC-2002.  
XX  
PF 12-FEB-2002; 2002US-00074956.  
XX  
PR 12-FEB-2001; 2001US-0268175P.  
XX  
PA (HEDL/) HEDLEY M L.  
XX  
PI Hedley ML;  
XX  
DR WPI; 2003-447327/42.  
XX  
PT Modulating immune responses in a mammal with a bladder disorder e.g.  
PT bladder cancer, by administering nucleic acids comprising un-methylated  
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
PT the mammal.  
XX  
PS Example 2; Page 9; 17pp; English.  
XX  
CC The invention describes a method of modulating an immune response in a  
CC mammal, comprising identifying a mammal that has or is at risk for having  
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated  
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;  
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
CC method is useful for modulating immune response in a mammal having a  
CC bladder disorder, where administration of (N1) results in an amelioration  
CC of one or more symptoms of the disorder. Preferably, the method is useful  
CC for modulating immune response in a mammal having bladder cancer (where  
CC administration of (N1) results in a decrease in tumour size or activity),  
CC or for modulating immune response in a mammal having interstitial  
CC cystitis (where administration of (N1) results in a modulation of the  
CC immune response from Th2 response to a Th1 response). The method is also  
CC useful for modulating immune response in a mammal having bladder disorder

CC that is characterised by inflammation which is associated with symptoms  
CC of interstitial cystitis or associated with a disruption of the integrity  
CC of the bladder lining. This is the amino acid sequence of a human serum  
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the  
CC treatment of bladder disorders  
XX  
SQ Sequence 241 AA;  
Query Match 100.0%; Score 47; DB 7; Length 241;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKQEPERNE 9  
Db 116 AKQEPERNE 124  
|||||  
RESULT 14  
AAO16985  
ID AAO16985 standard; protein; 242 AA.  
XX  
AC AAO16985;  
XX  
DT 29-MAY-2002 (first entry)  
XX  
DE Alpha-MSH construct protein fragment SEQ ID NO: 70.  
XX  
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritis;  
KW antidiabetic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX  
OS Unidentified.  
XX  
PN WO200206316-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 16-JUL-2001; 2001WO-US022263.  
XX  
PR 14-JUL-2000; 2000US-0218381P.  
PR 18-AUG-2000; 2000US-0226382P.  
PR 06-OCT-2000; 2000US-0238380P.  
PR 29-DEC-2000; 2000US-0258764P.  
PR 14-JUN-2001; 2001US-0298317P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.  
XX  
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory  
PT or autoimmune disorders.  
XX  
PS Disclosure; Page 5; 89pp; English.  
XX  
CC The present invention relates to a nucleic acid comprising a sequence  
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
CC hormone (MSH) concatamer. The sequences are useful for treating an  
CC individual suffering from, or at risk of, a disorder of the immune system  
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
CC present sequence is a protein described in the exemplification of the  
CC invention  
XX  
SQ Sequence 242 AA;

Db 116 AKQEPERNE 124

Search completed: April 19, 2004, 11:51:17

Job time : 5.29363 secs

Query Match 100.0%; Score 47; DB 5; Length 242;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

Db 116 AKQEPERNE 124

RESULT 15

AAO16986

ID AAO16986 standard; protein; 244 AA.

XX AAO16986;

AC AAO16986;

XX 29-MAY-2002 (first entry)

DT 29-MAY-2002 (first entry)

XX Alpha-MSH construct protein fragment SEQ ID NO: 71.

DE Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;

XX Alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;

KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;

KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;

KW antidiabetic; antibacterial; dermatological; antipsoriatic;

KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;

KW diabetes; uveitis; coeliac disease.

XX Unidentified.

OS WO200206316-A2.

PN 24-JAN-2002.

XX 16-JUL-2001; 2001WO-US022263.

PF 14-JUL-2000; 2000US-0218381P.

XX 18-AUG-2000; 2000US-0228382P.

PR 06-OCT-2000; 2000US-0238380P.

PR 29-DEC-2000; 2000US-0258764P.

PR 14-JUN-2001; 2001US-0298317P.

XX (ZYCO-) ZYCOS INC.

PA Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;

XX WPI; 2002-195801/25.

DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte

PT stimulating hormone concatamer or its analog, for treating inflammatory

PT or autoimmune disorders.

XX Disclosure; Page 5; 89pp; English.

PS The present invention relates to a nucleic acid comprising a sequence

XX encoding a fusion polypeptide having an alpha-melanocyte stimulating

CC hormone (MSH) concatamer. The sequences are useful for treating an

CC individual suffering from, or at risk of, a disorder of the immune system

CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid

CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact

CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,

CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The

CC present sequence is a protein described in the exemplification of the

CC invention

XX Sequence 244 AA;

SQ

Query Match 100.0%; Score 47; DB 5; Length 244;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

Db 116 AKQEPERNE 124

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_92\_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09E\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 47    | 100.0       | 195    | 13    | US-10-074-956-24  |
| 2          | 47    | 100.0       | 241    | 13    | US-10-074-956-27  |
| 3          | 47    | 100.0       | 268    | 13    | US-10-074-956-28  |
| 4          | 47    | 100.0       | 585    | 9     | US-09-929-552-2   |
| 5          | 47    | 100.0       | 585    | 10    | US-09-932-613-445 |
| 6          | 47    | 100.0       | 585    | 10    | US-09-984-010-26  |
| 7          | 47    | 100.0       | 585    | 10    | US-09-833-041-18  |
| 8          | 47    | 100.0       | 585    | 10    | US-09-833-117-18  |
| 9          | 47    | 100.0       | 585    | 10    | US-09-932-322-445 |
| 10         | 47    | 100.0       | 585    | 10    | US-09-832-501-18  |
| 11         | 47    | 100.0       | 585    | 11    | US-09-833-118-18  |
| 12         | 47    | 100.0       | 585    | 11    | US-09-833-245-18  |
| 13         | 47    | 100.0       | 585    | 12    | US-10-424-999-11  |
| 14         | 47    | 100.0       | 585    | 12    | US-10-425-000-31  |
| 15         | 47    | 100.0       | 585    | 12    | US-10-433-108-34  |

|    |    |       |     |    |                   |                    |
|----|----|-------|-----|----|-------------------|--------------------|
| 16 | 47 | 100.0 | 585 | 13 | US-10-153-064-5   | Sequence 5, Appli  |
| 17 | 47 | 100.0 | 585 | 14 | US-10-153-604A-5  | Sequence 5, Appli  |
| 18 | 47 | 100.0 | 585 | 14 | US-10-319-263-1   | Sequence 1, Appli  |
| 19 | 47 | 100.0 | 585 | 14 | US-10-319-263-2   | Sequence 2, Appli  |
| 20 | 47 | 100.0 | 585 | 14 | US-10-414-469-1   | Sequence 1, Appli  |
| 21 | 47 | 100.0 | 585 | 14 | US-10-414-469-2   | Sequence 2, Appli  |
| 22 | 47 | 100.0 | 585 | 14 | US-10-413-831-1   | Sequence 1, Appli  |
| 23 | 47 | 100.0 | 585 | 14 | US-10-413-831-2   | Sequence 2, Appli  |
| 24 | 47 | 100.0 | 585 | 15 | US-10-413-832-1   | Sequence 1, Appli  |
| 25 | 47 | 100.0 | 585 | 15 | US-10-413-832-2   | Sequence 2, Appli  |
| 26 | 47 | 100.0 | 585 | 15 | US-10-414-386-1   | Sequence 1, Appli  |
| 27 | 47 | 100.0 | 585 | 15 | US-10-414-386-2   | Sequence 2, Appli  |
| 28 | 47 | 100.0 | 585 | 15 | US-10-233-675A-11 | Sequence 11, Appli |
| 29 | 47 | 100.0 | 585 | 15 | US-10-462-262-26  | Sequence 26, Appli |
| 30 | 47 | 100.0 | 604 | 10 | US-09-984-010-7   | Sequence 7, Appli  |
| 31 | 47 | 100.0 | 608 | 14 | US-10-165-603-24  | Sequence 24, Appli |
| 32 | 47 | 100.0 | 608 | 14 | US-10-165-603-25  | Sequence 25, Appli |
| 33 | 47 | 100.0 | 608 | 14 | US-10-316-253-42  | Sequence 42, Appli |
| 34 | 47 | 100.0 | 609 | 10 | US-09-919-039-370 | Sequence 370, App  |
| 35 | 47 | 100.0 | 609 | 12 | US-10-609-346-12  | Sequence 12, Appli |
| 36 | 47 | 100.0 | 609 | 13 | US-10-153-064-7   | Sequence 7, Appli  |
| 37 | 47 | 100.0 | 609 | 14 | US-10-153-064A-7  | Sequence 7, Appli  |
| 38 | 47 | 100.0 | 609 | 14 | US-10-365-623-23  | Sequence 23, Appli |
| 39 | 47 | 100.0 | 610 | 9  | US-09-984-186-2   | Sequence 2, Appli  |
| 40 | 47 | 100.0 | 610 | 14 | US-10-237-667-2   | Sequence 2, Appli  |
| 41 | 47 | 100.0 | 610 | 14 | US-10-237-708-2   | Sequence 2, Appli  |
| 42 | 47 | 100.0 | 610 | 14 | US-10-237-866-2   | Sequence 2, Appli  |
| 43 | 47 | 100.0 | 610 | 14 | US-10-237-871-2   | Sequence 2, Appli  |
| 44 | 47 | 100.0 | 610 | 14 | US-10-237-624-2   | Sequence 2, Appli  |
| 45 | 47 | 100.0 | 616 | 12 | US-10-433-108-13  | Sequence 13, Appli |

ALIGNMENTS

RESULT 1  
US-10-074-956-24  
; Sequence 24, Application US/10074956  
; Publication No. US20020193332A1  
; GENERAL INFORMATION:  
; APPLICANT: Hedley, Mary Lynne  
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS  
; FILE REFERENCE: 08191-022001  
; CURRENT APPLICATION NUMBER: US/10/074,956  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: 60/268,175  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-074-956-24

Query Match 100.0%; Score 47; DB 13; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9  
Db 92 AKQEPERNE 100

RESULT 2  
US-10-074-956-27  
; Sequence 27, Application US/10074956  
; Publication No. US20020193332A1  
; GENERAL INFORMATION:  
; APPLICANT: Hedley, Mary Lynne  
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS  
; FILE REFERENCE: 08191-022001  
; CURRENT APPLICATION NUMBER: US/10/074,956

```
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-27

Query Match          100.0%; Score 47; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      116 AKQEPERNE 124

RESULT 3
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08:91-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-28

Query Match          100.0%; Score 47; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      116 AKQEPERNE 124

RESULT 4
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match          100.0%; Score 47; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 5
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091585A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match          100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 6
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
```

```
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/984,010
/ FILING DATE: 21-May-2002
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 09/091,873
/ FILING DATE: 25-JUN-1998
/ APPLICATION NUMBER: PCT/GB96/03164
/ FILING DATE: 19-DEC-1996
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 585 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEetical: NO
/ ANTI-SENSE: NO
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 7
US-09-833-041-18
/ Sequence 18, Application US/09833041
/ Publication No. US20030125247A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, Craig A.
/ APPLICANT: Haseltine, William A.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PF545
/ CURRENT APPLICATION NUMBER: US/09/833,041
/ CURRENT FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229,358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256,931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199,384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 585
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 8
US-09-833-117-18
```

```
/ Sequence 18, Application US/09833117
/ Publication No. US20030171267A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, Craig A.
/ APPLICANT: Sadechi, Homa
/ APPLICANT: Prior, Christopher P.
/ APPLICANT: Turner, Andrew J.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PF543
/ CURRENT APPLICATION NUMBER: US/09/833,117
/ CURRENT FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229,358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256,931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199,384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 585
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 9
US-09-932-322-445
/ Sequence 445, Application US/09932322
/ Publication No. US20030194743A1
/ GENERAL INFORMATION:
/ APPLICANT: Dyax Corp.
/ APPLICANT: Beltzer, James P.
/ APPLICANT: Potter, M. Daniel
/ APPLICANT: Fleming, Tony J.
/ APPLICANT: Ladner, Robert Charles
/ TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
/ FILE REFERENCE: Dyx-018.1 PCT; DYX-018.1 US
/ CURRENT APPLICATION NUMBER: US/09/932,322
/ CURRENT FILING DATE: 2001-08-17
/ NUMBER OF SEQ ID NOS: 458
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 445
/ LENGTH: 585
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 10
US-09-832-501-18
/ Sequence 18, Application US/09832501
/ Publication No. US20030199043A1
/ GENERAL INFORMATION:
/ APPLICANT: Ballance, David J.
/ APPLICANT: Sleep, Darrell
/ APPLICANT: Turner, Andrew J.
```

```
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS42
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 11
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS44
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 12
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 13
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for:
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 14
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Xringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31
```

```
Query Match      100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AKQEPERNE 9
      |||||
Db      92 AKQEPERNE 100
```

## RESULT 15

```
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34
```

```
Query Match      100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AKQEPERNE 9
      |||||
Db      92 AKQEPERNE 100
```

```
Search completed: April 19, 2004, 12:54:58
Job time : 3.78947 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 0.919668 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | ID  | Description          |
|------------|-------|-------------|-----|----------------------|
| 1          | 33    | 100.0       | 1   | US-08-448-196A-1     |
| 2          | 33    | 100.0       | 2   | US-08-709-923-3      |
| 3          | 33    | 100.0       | 4   | US-09-252-991A-27640 |
| 4          | 33    | 100.0       | 582 | US-08-134-638-1      |
| 5          | 33    | 100.0       | 583 | US-08-448-196A-4     |
| 6          | 33    | 100.0       | 583 | US-08-448-196A-5     |
| 7          | 33    | 100.0       | 583 | US-08-448-196A-6     |
| 8          | 33    | 100.0       | 584 | US-08-448-196A-7     |
| 9          | 33    | 100.0       | 585 | US-08-153-799-14     |
| 10         | 33    | 100.0       | 585 | US-08-448-196A-3     |
| 11         | 33    | 100.0       | 585 | US-08-984-176-1      |
| 12         | 33    | 100.0       | 585 | US-08-702-572-2      |
| 13         | 33    | 100.0       | 585 | US-08-769-746-2      |
| 14         | 33    | 100.0       | 585 | US-10-153-064-5      |
| 15         | 33    | 100.0       | 609 | US-08-222-619-3      |
| 16         | 33    | 100.0       | 609 | US-08-433-037-4      |
| 17         | 33    | 100.0       | 609 | US-08-897-956A-2     |
| 18         | 33    | 100.0       | 609 | US-10-153-064-7      |
| 19         | 33    | 100.0       | 609 | US-09-976-594-73     |
| 20         | 33    | 100.0       | 609 | PCT-US95-04075-3     |
| 21         | 33    | 100.0       | 610 | US-08-797-689-2      |
| 22         | 33    | 100.0       | 610 | US-09-984-186-2      |
| 23         | 33    | 100.0       | 651 | US-10-153-064-133    |
| 24         | 33    | 100.0       | 652 | US-10-153-064-99     |
| 25         | 33    | 100.0       | 652 | US-10-153-064-96     |
| 26         | 33    | 100.0       | 652 | US-10-153-064-105    |
| 27         | 33    | 100.0       | 652 | US-10-153-064-132    |

28 33 100.0 653 4 US-10-153-064-131 Sequence 131, Appl  
29 33 100.0 656 4 US-10-153-064-130 Sequence 130, Appl  
30 33 100.0 660 4 US-10-153-064-90 Sequence 90, Appl  
31 33 100.0 660 4 US-10-153-064-93 Sequence 93, Appl  
32 33 100.0 666 4 US-09-489-039A-12388 Sequence 12388 A  
33 33 100.0 668 4 US-10-153-064-102 Sequence 102, Appl  
34 33 100.0 676 4 US-10-153-064-95 Sequence 95, Appl  
35 33 100.0 676 4 US-10-153-064-98 Sequence 98, Appl  
36 33 100.0 676 4 US-10-153-064-104 Sequence 104, Appl  
37 33 100.0 676 4 US-10-153-064-127 Sequence 127, Appl  
38 33 100.0 676 4 US-10-153-064-129 Sequence 129, Appl  
39 33 100.0 677 4 US-10-153-064-125 Sequence 125, Appl  
40 33 100.0 680 4 US-10-153-064-123 Sequence 123, Appl  
41 33 100.0 684 4 US-10-153-064-92 Sequence 92, Appl  
42 33 100.0 692 4 US-10-153-064-101 Sequence 101, Appl  
43 33 100.0 783 1 US-08-256-938-2 Sequence 2, Appl  
44 33 100.0 787 1 US-08-256-938-4 Sequence 4, Appl  
45 33 100.0 787 2 US-08-797-689-16 Sequence 16, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-448-196A-1

; Sequence 1, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TITLE OF INVENTION: RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSER: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: BROAD JR., ROBERT L.

; REGISTRATION NUMBER: 18,757

; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 205-544-0021

; TELEFAX: 205-544-0258

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 109 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-08-448-196A-1

Query Match 100.0%; Score 33; DB 1; Length 109;

Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HGDLE 6

Db 57 HGDLE 62



RESULT 2  
US-08-709-923-3  
; Sequence 3, Application US/08709923  
; Patent No. 5831055  
; GENERAL INFORMATION:  
; APPLICANT: BIENKOWSKI, MICHAEL J.  
; TITLE OF INVENTION: NOVEL KIDNEY ATP-DEPENDENT POTASSIUM  
; CHANNELS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pharmacia & Upjohn Company  
; STREET: 301 Henrietta Street  
; CITY: Kalamazoo  
; STATE: MI  
; COUNTRY: USA  
; ZIP: 49001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709, 923  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wootton, Thomas A.  
; REGISTRATION NUMBER: 35,004  
; REFERENCE/DOCKET NUMBER: 6001.N CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 616-833-7914  
; TELEFAX: 616-833-6897  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-709-923-3

Query Match 100.0%; Score 33; DB 2; Length 380;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
Db 93 HGDLLLE 98

RESULT 3  
US-09-252-991A-27640  
; Sequence 27640, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27640  
; LENGTH: 547

; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27640  
Query Match 100.0%; Score 33; DB 4; Length 547;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
|||||  
Db 488 HGDLLLE 493  
RESULT 4  
US-08-134-638-1  
; Sequence 1, Application US/08134638  
; Patent No. 5473050  
; GENERAL INFORMATION:  
; APPLICANT: Strand, Frederick T  
; TITLE OF INVENTION: Denatured Bovine Serum Albumin Milk  
; TITLE OF INVENTION: Products and Method Therefor  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Frederick T. Strand  
; STREET: P.O. Box 64321  
; CITY: Phoenix  
; STATE: Arizona  
; COUNTRY: USA  
; ZIP: 85082-4321  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb storage  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: MS-DOS 5.0  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/134,638  
; FILING DATE: 10/12/93  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: N/A  
; FILING DATE: N/A  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weiss, Harry M  
; REGISTRATION NUMBER: 19,497  
; REFERENCE/DOCKET NUMBER: 1795P1423  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (602) 994-8888  
; TELEFAX: (602) 947-2683  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 582  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-134-638-1

Query Match 100.0%; Score 33; DB 1; Length 582;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6  
|||||  
Db 245 HGDLLLE 250

RESULT 5  
US-08-448-196A-4  
; Sequence 4, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

;; TITLE OF INVENTION: RELATED PROTEINS  
;; NUMBER OF SEQUENCES: 9  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NASA  
;; STREET: MARSHALL SPACE FLIGHT CENTER  
;; CITY: HUNTSVILLE  
;; STATE: ALABAMA  
;; COUNTRY: USA  
;; ZIP: 35812  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent in Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 23-MAY-1995  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BROAD JR., ROBERT L.  
;; REGISTRATION NUMBER: 18,757  
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
;; TELEPHONE: 205-544-0021  
;; TELEFAX: 205-544-0258  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 583 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: N-terminal  
;; APPLICATION NUMBER: US/08/448,196A  
;; FILING DATE: 23-MAY-1995  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BROAD JR., ROBERT L.  
;; REGISTRATION NUMBER: 18,757  
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
;; TELEPHONE: 205-544-0021  
;; TELEFAX: 205-544-0258  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 583 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: N-terminal  
;; APPLICATION NUMBER: US/08/448,196A-4  
US-08-448-196A-4  
Query Match 100.0%; Score 33; DB 1; Length 583;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
Db 246 HGDLLLE 251  
RESULT 6  
US-08-448-196A-5  
; Sequence 5, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; APPLICATION NUMBER: US/08/448,196A-6  
US-08-448-196A-6  
Query Match 100.0%; Score 33; DB 1; Length 583;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;; REGISTRATION NUMBER: 18,757  
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 205-544-0021  
;; TELEFAX: 205-544-0258  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 583 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: N-terminal  
;; APPLICATION NUMBER: US/08/448-196A-5  
US-08-448-196A-5  
Query Match 100.0%; Score 33; DB 1; Length 583;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HGDLLLE 6  
Db 246 HGDLLLE 251  
RESULT 7  
US-08-448-196A-6  
; Sequence 6, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; APPLICATION NUMBER: US/08/448-196A-6  
US-08-448-196A-6  
Query Match 100.0%; Score 33; DB 1; Length 583;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
Db 246 HGDLL 251

RESULT 8  
US-08-448-196A-7  
; Sequence 7, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; CITY: MARSHALL SPACE FLIGHT CENTER  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 584 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-448-196A-7

Query Match 100.0%; Score 33; DB 1; Length 584;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
Db 247 HGDLL 252

RESULT 9  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA

QY 1 HGDLL 6  
Db 247 HGDLL 252

RESULT 10  
US-08-448-196A-3  
; Sequence 3, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER

US-08-153-799-14

Query Match 100.0%; Score 33; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6  
Db 247 HGDLL 252

RESULT 10  
US-08-448-196A-3  
; Sequence 3, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER

```

; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-3

Query Match 100.0%; Score 33; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252

RESULT 11
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 33; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252

RESULT 12
US-08-702-572-2
; Sequence 2, Application US/08702572
```

```

; Patent No. 5963386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 33; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252

RESULT 13
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-746-2

Query Match 100.0%; Score 33; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLL 6
Db 247 HGDLLL 252

RESULT 14
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-153-064-5

Query Match 100.0%; Score 33; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLL 6
Db 247 HGDLLL 252

RESULT 15
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California

; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-746-2

Query Match 100.0%; Score 33; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLL 6
Db 271 HGDLLL 276

Search completed: April 19, 2004, 12:05:20
Job time : 1.91967 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.4626 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_266\_277

Perfect score: 57

Sequence: 1 ENQDSISSKLKE 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 57    | 100.0       | 600    | 2 A47391 | serum albumin prec  |
| 2          | 57    | 100.0       | 609    | 1 ABRUS  | serum albumin prec  |
| 3          | 54    | 94.7        | 265    | 2 I45986 | albumin - dog (fra  |
| 4          | 54    | 94.7        | 608    | 2 S57632 | serum albumin prec  |
| 5          | 51    | 89.5        | 605    | 1 ABFGS  | serum albumin prec  |
| 6          | 51    | 89.5        | 607    | 1 ABRBS  | serum albumin prec  |
| 7          | 45    | 78.9        | 423    | 2 S41122 | mannose-6-phosphat  |
| 8          | 44    | 77.2        | 599    | 1 A54906 | afamin precursor -  |
| 9          | 44    | 77.2        | 607    | 1 ABRHS  | serum albumin prec  |
| 10         | 43    | 75.4        | 607    | 1 ABRGS  | serum albumin prec  |
| 11         | 40    | 70.2        | 609    | 2 JCS838 | albumin - Mongolia  |
| 12         | 40    | 70.2        | 2469   | 2 H36812 | hypothetical prote  |
| 13         | 39    | 68.4        | 608    | 2 A53195 | afamin precursor -  |
| 14         | 38    | 66.7        | 238    | 2 T32033 | hypothetical prote  |
| 15         | 38    | 66.7        | 1091   | 2 T34107 | hypothetical prote  |
| 16         | 37    | 64.9        | 416    | 2 P68640 | hippurate hydratase |
| 17         | 37    | 64.9        | 453    | 2 A05139 | serum albumin - mo  |
| 18         | 37    | 64.9        | 495    | 2 F82123 | probable flagellar  |
| 19         | 37    | 64.9        | 535    | 2 S31097 | cold acclimation p  |
| 20         | 37    | 64.9        | 556    | 2 B28882 | conserved hypothet  |
| 21         | 37    | 64.9        | 563    | 2 AH2975 | hypothetical prote  |
| 22         | 37    | 64.9        | 563    | 2 C98307 | opha protein limpo  |
| 23         | 37    | 64.9        | 581    | 2 P84857 | hypothetical prote  |
| 24         | 37    | 64.9        | 608    | 1 ABRIS  | serum albumin prec  |
| 25         | 36    | 63.2        | 176    | 2 F81346 | heat shock protein  |
| 26         | 36    | 63.2        | 208    | 2 B30305 | submandibular glan  |
| 27         | 36    | 63.2        | 282    | 2 T24693 | hypothetical prote  |
| 28         | 36    | 63.2        | 298    | 2 G75140 | hypothetical prote  |
| 29         | 36    | 63.2        | 419    | 2 T44276 | transducer protein  |

|    |    |      |      |          |                    |
|----|----|------|------|----------|--------------------|
| 30 | 36 | 63.2 | 420  | 2 C84298 | Htr12 transducer [ |
| 31 | 36 | 63.2 | 575  | 2 C88924 | protein R02C2.4 [i |
| 32 | 36 | 63.2 | 1011 | 2 S65668 | preprotein translo |
| 33 | 36 | 63.2 | 1037 | 2 D96786 | protein F10A5.15 [ |
| 34 | 36 | 63.2 | 1170 | 2 S30010 | probable finger pr |
| 35 | 36 | 63.2 | 2471 | 2 T42977 | large tegument pro |
| 36 | 35 | 61.4 | 67   | 2 G72392 | heavy metal bindin |
| 37 | 35 | 61.4 | 79   | 2 B9761  | hypothetical prote |
| 38 | 35 | 61.4 | 90   | 2 T44137 | hypothetical prote |
| 39 | 35 | 61.4 | 248  | 2 A69173 | conserved hypothet |
| 40 | 35 | 61.4 | 264  | 2 T33614 | hypothetical prote |
| 41 | 35 | 61.4 | 296  | 1 G64453 | conserved hypothet |
| 42 | 35 | 61.4 | 314  | 2 AC1484 | B. subtilis YgfA p |
| 43 | 35 | 61.4 | 314  | 2 A11123 | B. subtilis YgfA p |
| 44 | 35 | 61.4 | 317  | 2 A48443 | merozoite 44K vari |
| 45 | 35 | 61.4 | 416  | 2 S52075 | farnesyl-diphospha |

ALIGNMENTS

RESULT 1

A47391  
serum albumin precursor - rhesus macaque  
C;Species: Macaca mulatta (rhesus macaque)  
C;Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C;Accession: A47391  
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir  
A;Reference number: A47391; MUID:93211971; PMID:8460152  
A;Contents: B/B homozygote  
A;Accession: A47391  
A;Status: preliminary  
A;Molecule type: mRNA; protein  
A;Residues: 1-600 <WAT>  
A;Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295  
A;Experimental source: liver  
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIIP:128281)  
C;Superfamily: serum albumin; serum albumin repeat homology  
F;21-194/Domain: serum albumin repeat homology <SA1>  
F;213-386/Domain: serum albumin repeat homology <SA2>  
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 57; DB 2; Length 600;  
Best Local Similarity 100.0%; Pred. No. 0.009;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ENQDSISSKLKE 12  
|||  
Db 282 ENQDSISSKLKE 293

RESULT 2

ABRUS  
serum albumin precursor [validated] - human  
N;Alternate names: preproalbumin  
N;Contains: kinetensin  
C;Species: Homo sapiens (man)  
C;Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text\_change 17-Mar-2000  
C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3  
R;Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu  
Nucleic Acids Res. 9, 6103-6114, 1981  
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col  
A;Reference number: A93743; MUID:82081882; PMID:6171778  
A;Accession: A93743  
A;Molecule type: mRNA  
A;Residues: 1-419, 'K', 421-609 <LAW>  
A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:CAA2  
R;Dugaiczky, A.; Law, S.W.; Dennison, O.E  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A;Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936  
 A:Molecule type: mRNA  
 A:Residues: 1-120,'G',122-609 <DUG>  
 A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
 R:Umano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
 J. Biol. Chem. 261, 3244-3251, 1986  
 A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
 A:Reference number: 139427; MUID:86140099; PMID:2419329  
 A:Accession: I39427  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-26 <RA>  
 A:Cross-references: GB:M13075; NID:G178330; PIDN:AA51688.1; PID:G553173  
 R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
 A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
 A:Reference number: 159286; MUID:94181575; PMID:8134387  
 A:Accession: I59286  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 282-290, 'KSRPDLQ' <WAT>  
 A:Cross-references: GB:S69192; NID:G546032; PIDN:AA30282.1; PID:G546033  
 A:Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia  
 R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, F.W.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
 A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
 A:Reference number: 159313; MUID:94294404; PMID:8022807  
 A:Accession: I59313  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 589-590, 'ALPRPVKNLLQVLP' <MAD>  
 A:Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232  
 A:Note: This frame-shift variant is designated albumin Bazzano; four additional variants  
 R:Menaya, J.; Parrilla, R.; Ayuso, M.S.  
 submitted to the EMBL Data Library, March 1995  
 A:Reference number: G08292  
 A:Accession: G01747  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-120,'G',122-455 <MEN>  
 A:Cross-references: EMBL:U22361; NID:G763428; PIDN:AA64922.1; PID:G763431  
 R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
 Biochem. J. 308, 321-325, 1995  
 A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*  
 A:Reference number: S55314; MUID:95275251; PMID:7755581  
 A:Accession: S55314  
 A:Molecule type: protein  
 A:Residues: 19-27 <LED>  
 R:Meloun, B.; Moravek, L.; Kostka, V.  
 FEBS Lett. 58, 134-137, 1975  
 A:Title: Complete amino acid sequence of human serum albumin.  
 A:Reference number: A91420; MUID:76187907; PMID:1225573  
 A:Accession: A91420  
 A:Molecule type: protein  
 A:Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-  
 R:Roehr, U.; Spitteller, G.; Tripiet, D.  
 Justus Liebig Ann. Chem. 9, 881-884, 1988  
 A:Title: Isolation and structure elucidation of middle-molecular weight peptides from ur  
 A:Reference number: S06422  
 A:Accession: S06422  
 A:Molecule type: protein  
 A:Residues: 25-48 <ROE>  
 R:Pinch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
 Arch. Biochem. Biophys. 305, 595-599, 1993  
 A:Title: Mass spectrometric identification of modifications to human serum albumin treat  
 A:Reference number: S36882; MUID:93384321; PMID:8373198  
 A:Accession: S36882  
 A:Molecule type: protein  
 A:Residues: 45-67,'141-160',311-337,'469-490',570-581 <PIN>  
 R:Kausler, E.; Spitteller, G.  
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
 A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A:Reference number: S17599; MUID:92126241; PMID:1772598  
 A:Accession: S17599  
 A:Molecule type: protein  
 A:Residues: 25-54,'354-357',431-447 <KAU>  
 A:Note: 49-Leu was also found  
 R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
 J. Immunol. 143, 1680-1684, 1989  
 A:Title: Structures of histamine-releasing peptides formed by the action of acid protea  
 A:Reference number: A45800; MUID:89341406; PMID:2474609  
 A:Accession: A45800  
 A:Molecule type: protein  
 A:Residues: 166-173 <CAR>  
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W.  
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
 A:Title: The amino acid sequence of kinitensin, a novel peptide isolated from pepsin-tr  
 A:Reference number: A03239; MUID:86242180; PMID:3087352  
 A:Accession: A03239  
 A:Molecule type: protein  
 A:Residues: 166-173,'L', <MOG>  
 R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins,  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.  
 A:Reference number: A38255; MUID:91062352; PMID:2247440  
 A:Accession: C38255  
 A:Molecule type: protein  
 A:Residues: 76-111 <GALL>  
 A:Accession: B38255  
 A:Molecule type: protein  
 A:Residues: 82-105,'K',107-110 <GAL2>  
 A:Note: this variant is designated albumin Vibo Valentia  
 A:Accession: A38255  
 A:Molecule type: protein  
 A:Residues: 76-83,'K',85-106 <GAL3>  
 A:Note: this variant is designated albumin Torino  
 R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
 Eur. J. Biochem. 214, 437-444, 1993  
 A:Title: The structural characterization and bilirubin-binding properties of albumin He  
 A:Reference number: S33298; MUID:93292504; PMID:8513793  
 A:Accession: S33298  
 A:Molecule type: protein  
 A:Residues: 255-263,'E',265-281 <MIN1>  
 A:Note: this variant is designated albumin Herborn  
 R:Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta  
 Biochim. Biophys. Acta 1119, 232-238, 1992  
 A:Title: Two albumins with identical electrophoretic mobility are produced by diffe  
 A:Reference number: S21078; MUID:92190239; PMID:1347703  
 A:Accession: S21078  
 A:Molecule type: protein  
 A:Residues: 354-356,'K',358-378 <MIN2>  
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported  
 R:He, X.M.; Carter, D.C.  
 Nature 358, 209-215, 1992  
 A:Title: Atomic structure and chemistry of human serum albumin.  
 A:Reference number: A46756; MUID:92334427; PMID:1630489  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.  
 in: The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40  
 A:Reference number: A94442  
 A:Contents: annotation; three-dimensional structure and disulfide bonds  
 R:Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A:Title: Disulfide bonds in human serum albumin.  
 A:Reference number: A90930  
 A:Contents: annotation; disulfide bonds  
 R:Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A:Reference number: A90299; MUID:78186630; PMID:656055  
 A:Contents: annotation; bilirubin-binding site  
 R:Peterson, T.; Reed, R.G.  
 in: Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20  
 A:Title: Serum albumin: conformation and active sites.  
 A:Reference number: A94408

A;Contents: annotation; binding sites  
R;Harper, M.E.; Dugaiczky, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
A;Reference number: A90028; MUID:8327982; PMID:6192711  
A;Contents: annotation; gene position  
R;Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
A;Reference number: A46755; MUID:76257808; PMID:955075  
A;Contents: annotation  
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)  
R;Bohney, J.P.; Fonda, M.L.; Feildhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A;Title: Identification of lys(190) as the primary binding site for pyridoxal 5'-phosphate  
A;Reference number: A56294; MUID:92183881; PMID:1544460  
A;Contents: annotation  
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in pyridoxal phosphate  
A;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak C;Comment: A large number of variants of human serum albumin have been described.  
C;Genetics:  
A;Gene: GDB:ALB  
A;Cross-references: GDB:118990; OMIM:103600  
A;Map position: 4q11-4q13  
C;Superfamily: serum albumin; serum albumin repeat homology  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-609/Product: serum albumin #status experimental <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;166-174/Product: kinetensin #status experimental <KIP>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,148-193,192-203,224-270,269-277,289-303,302-313,340-385,394-393,4  
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 57; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 0.0092;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12  
|||||:|||||  
DB 290 ENQDSISSKLKE 301

RESULT 3  
ABFGS  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-605 <WEI>  
A;Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798  
C;Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 20-Aug-1999  
C;Accession: I46986  
R;Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner  
J. Allergy Clin. Immunol. 93, 614-627, 1994  
A;Title: Molecular characterization of dog albumin as a cross-reactive allergen.  
A;Reference number: I46986; MUID:94204492; PMID:7512102  
A;Accession: I46986  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-265 <SPI>  
A;Cross-references: GB:S72946; NID:G633937; PIDN:AA330434.1; PID:G633938  
C;Superfamily: serum albumin; serum albumin repeat homology  
F;7-180/Domain: serum albumin repeat homology <SA2>

Query Match 94.7%; Score 54; DB 2; Length 265;  
Best Local Similarity 91.7%; Pred. No. 0.014;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12  
|||||:|||||  
DB 76 ENQDSISSKLKE 87

## RESULT 4

S57632  
serum albumin precursor - cat  
C;Species: Felis silvestris catus (domestic cat)  
C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C;Accession: J4660; S57632  
R;Hilger, C.; Grigioni, F.; Hentges, F.  
Gene 169, 295-296, 1996  
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A;Reference number: J4660; MUID:96194924; PMID:8647469  
A;Accession: J4660  
A;Molecule type: mRNA  
A;Residues: 1-608 <HI2>  
A;Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
A;Experimental source: liver  
C;Comment: This protein is the major protein component in plasma. It functions as a multi-chain protein with 35 conserved cysteine residues.  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: liver; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRP>  
F;25-608/Product: serum albumin #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 94.7%; Score 54; DB 2; Length 608;  
Best Local Similarity 91.7%; Pred. No. 0.034;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12  
|||||:|||||  
DB 290 ENQDSISSKLKE 301

RESULT 5  
ABFGS  
serum albumin precursor - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: S01382; A61006  
R;Weinstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 9045, 1988  
A;Title: Nucleotide sequence of porcine liver albumin.  
A;Reference number: S01382; MUID:89016582; PMID:3174440  
A;Accession: S01382  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-605 <WEI>  
A;Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798  
R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral  
A;Reference number: A61006; MUID:89269769; PMID:2728927  
A;Accession: A61006  
A;Molecule type: protein  
A;Residues: 23-51, 'X', 53-54; 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XV', 155 <LIM>  
A;Experimental source: dental enamel  
A;Note: albumin and other serum proteins are also found in bone  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, steroid hormones (weak bonds with these hormones promote their transfer across the membrane).  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-16/Domain: signal sequence (fragment) #status predicted <PRO>  
F;17-22/Domain: propeptide #status predicted <PRP>  
F;23-605/Product: serum albumin #status predicted <MAT>  
F;27-199/Domain: serum albumin repeat homology <SA1>  
F;218-391/Domain: serum albumin repeat homology <SA2>  
F;410-589/Domain: serum albumin repeat homology <SA3>  
F;75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390, 4  
F;261/Binding site: bilirubin (Lys) #status predicted



Query Match 89.5%; Score 51; DB 1; Length 605;  
Best Local Similarity 83.3%; Pred. No. 0.13;  
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12  
|||||:|||||  
Db 287 ENQDITSLKKE 298

RESULT 6  
ABOS  
serum albumin precursor [validated] - bovine  
N/Alternate names: 67K protein; prealbumin  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 18-Aug-2000  
C/Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94  
R/Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.  
submitted to the EMBL Data Library, August 1991  
A/Description: Bovine serum albumin: cDNA sequence and expression.  
A/Reference number: A38885  
A/Accession: A38885  
A/Molecule type: mRNA  
A/Residues: 1-607 <HL>  
A/Cross-references: EMBL:W73215  
R/Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
A/Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
A/Reference number: A36401; MUID:91083649; PMID:2260975  
A/Accession: A36401  
A/Molecule type: protein  
A/Residues: 25-41, H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HR>  
R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979  
A/Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A/Reference number: A91258; MUID:80024278; PMID:488109  
A/Accession: A91258  
A/Molecule type: protein  
A/Residues: 1-32 <MG>  
R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A/Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing g  
A/Reference number: A60808; MUID:88267456; PMID:3389500  
A/Accession: B60808  
A/Molecule type: protein  
A/Residues: 25-41 <HSI>  
R/Strawich, E.; Glincher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A/Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is albu  
A/Reference number: S10780; MUID:90336641; PMID:2379503  
A/Accession: S10780  
A/Molecule type: protein  
A/Residues: 25-41, H', 43-57, 59-64 <STR>  
R/Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A/Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A/Reference number: A45800; MUID:89341406; PMID:2474609  
A/Accession: D45800  
A/Molecule type: protein  
A/Residues: 163-172 <CAR>  
R/Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A/Title: Structure of a biologically active neurotensin-related peptide obtained from pe  
A/Reference number: A26693; MUID:87194805; PMID:2437111  
A/Accession: A26693  
A/Molecule type: protein  
A/Residues: 165-172, 'L', <CA2>  
R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A/Title: Sequence of residues 400-403 of bovine serum albumin.  
A/Reference number: A90309; MUID:82023364; PMID:7283978  
A/Accession: A90309  
A/Molecule type: protein

A/Residues: 402-433 <REE>  
R/Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A/Title: Structure of bovine serum albumin.  
A/Reference number: A91458  
A/Accession: A91458  
A/Molecule type: protein  
A/Residues: 25-41, H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'I'  
R/Brown, J.R.  
submitted to the Atlas, April 1975  
A/Reference number: A94551  
A/Accession: A94551  
A/Molecule type: protein  
A/Residues: 190-195 <BR2>  
R/Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A/Reference number: A91457  
A/Contents: annotation; disulfide bonds  
R/Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A/Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
A/Reference number: S55232; MUID:95031935; PMID:7945219  
A/Accession: S55232  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 529-536; 569-572 <WER>  
C/Suprafamily: serum albumin; serum albumin repeat homology  
C/Keywords: carrier protein; copper binding; duplication; plasma  
F/1-18/Domain: signal sequence #status experimental <SIG>  
F/19-24/Domain: propeptide #status experimental <PRO>  
F/23-607/Product: serum albumin #status experimental <MPT>  
F/23-201/Domain: serum albumin repeat homology <SA1>  
F/220-393/Domain: serum albumin repeat homology <SA2>  
F/412-591/Domain: serum albumin repeat homology <SA3>  
F/21/Binding site: copper (His) #status predicted  
F/77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, '  
Query Match 89.5%; Score 51; DB 1; Length 607;  
Best Local Similarity 83.3%; Pred. No. 0.13;  
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12  
|||||:|||||  
Db 289 DNQDTISSKLKE 300

RESULT 7  
S41122  
mannose-6-phosphate isomerase (EC 5.3.1.8) - human  
N/Alternate names: phosphomannose isomerase  
C/Species: Homo sapiens (man)  
C/Date: 07-Sep-1994 #sequence revision 10-Nov-1995 #text\_change 08-Oct-1999  
C/Accession: S41122; S38666  
R/Proudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.  
Eur. J. Biochem. 219, 415-423, 1994  
A/Title: Purification, cDNA cloning and heterologous expression of human phosphomannose  
A/Reference number: S41122; MUID:94139717; PMID:8307007  
A/Accession: S41122  
A/Molecule type: mRNA  
A/Residues: 1-423 <PRO>  
A/Cross-references: EMBL:X76057; NID:9416016; PIDN:CAAS3657.1; PID:9416017  
C/Genetics:  
A/Gene: GDB:MPI  
A/Cross-references: GDB:119397; OMIM:154550  
A/Map position: 15q22-15qter  
C/Suprafamily: yeast mannose-6-phosphate isomerase  
C/Keywords: intramolecular oxidoreductase; isomerase

Query Match 78.9%; Score 45; DB 2; Length 423;  
Best Local Similarity 66.7%; Pred. No. 1.2;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12

Db 77 ENQDLSGSKVD 88  
|||||:|:|:|

RESULT 8  
ABSHS  
serum albumin precursor - human  
N/Alternate names: alpha-albumin  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 17-Mar-2000  
C/Accession: A54906; JG6143; S68554; S78082; I39425  
R/Lichenstein, H.S.; Lyons, D.E.; Wurfel, M.M.; Johnson, D.A.; McGinley, M.D.; Leidli, J.  
J. Biol. Chem. 269, 18149-18154, 1994  
A/Title: Afamin is a new member of the albumin, alpha-fetoprotein, and vitamin D-binding  
A/Reference number: A54906; MUID:94299534; PMID:7517938  
A/Accession: A54906  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-599 <LIC>  
A/Cross-references: GB:L32140; NID:G533885; PIDN:AAA21612.1; PID:G547402  
R/Mishio, H.; Dugaiczkyk, A.  
Proc. Natl. Acad. Sci. U.S.A. 93, 7557-7561, 1996  
A/Title: Complete structure of the human alpha-albumin gene, a new member of the serum  
A/Reference number: JG6143; MUID:96353855; PMID:8755513  
A/Accession: JG6143  
A/Molecule type: DNA  
A/Residues: 1-599 <NIS>  
A/Cross-references: GB:U51243; NID:G1418261; PIDN:AAC50720.1; PID:G1418262  
R/Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczkyk, A.  
J. Mol. Biol. 259, 113-119, 1996  
A/Title: Tandem arrangement of the human serum albumin multigene family in the sub-centr  
A/Reference number: S68554; MUID:96240683; PMID:8648639  
A/Accession: S68554  
A/Molecule type: DNA  
A/Residues: 1-29 <NIW>  
A/Cross-references: GB:U51243; NID:G1418261  
A/Note: neither the complete nucleic acid sequence nor the complete translation are show  
R/Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczkyk, A.  
submitted to the EMBL Data Library, March 1996  
A/Description: Tandem arrangement of the human serum albumin multigene family in the sub  
A/Reference number: S78082  
A/Accession: S78082  
A/Molecule type: DNA  
A/Residues: 1-599 <NIS>  
A/Cross-references: EMBL:U51243; NID:G1418261; PIDN:AAC50720.1; PID:G1418262  
R/Lallard, D.; Gilbert, S.; Lamontagne, A.; Hamel, D.; Belanger, L.  
Gene 153, 287-288, 1995  
A/Title: Identification of rat alpha-albumin and cDNA cloning of its human ortholog.  
A/Reference number: I39424; MUID:96180738; PMID:7875606  
A/Accession: I39425  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 105-207 <ALL>  
A/Cross-references: GB:L35497; NID:G530131; PIDN:AAA68198.1; PID:G857676  
C/Genetics:  
A/Genes: GDB:AFM  
A/Cross-references: GDB:376475; OMIM:104145  
A/Map position: 4q11-q13  
A/Introns: 30/1, 46/2, 90/3, 161/2, 205/3, 238/2, 281/3, 353/2, 397/3, 430/2, 474/3, 549  
C/Suprafamily: serum albumin; serum albumin repeat homology  
C/Keywords: extracellular protein; glycoprotein; plasma  
F/1-21/Domain: signal sequence #status predicted <SIG>  
F/22-599/Product: afamin #status predicted <MAT>  
F/29-202/Domain: serum albumin repeat homology <SA1>  
F/221-394/Domain: serum albumin repeat homology <SA2>  
F/413-590/Domain: serum albumin repeat homology <SA3>  
F/33,109,402,488/Binding site: carbohydrate (Asn) #status predicted

Query Match 77.2%; Score 44; DB 1; Length 599;  
Best Local Similarity 90.0%; Pred. No. 2.7;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDSISSKLE 12

Db 292 QDSISSKLE 301  
|||||:|:|:|

RESULT 9  
ABSHS  
serum albumin precursor - sheep  
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C/Accession: S06936  
R/Brown, W.M.; Dziągiewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A/Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A/Reference number: S06936; MUID:90098888; PMID:2602160  
A/Accession: S06936  
A/Molecule type: mRNA  
A/Residues: 1-607 <BRO>  
A/Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387  
A/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membr  
C/Suprafamily: serum albumin; serum albumin repeat homology  
C/Keywords: carrier protein; duplication; metal binding; plasma  
F/1-18/Domain: signal sequence #status predicted <SIG>  
F/19-24/Domain: propeptide #status predicted <PRO>  
F/25-607/Product: serum albumin #status predicted <MAT>  
F/29-201/Domain: serum albumin repeat homology <SA1>  
F/220-393/Domain: serum albumin repeat homology <SA2>  
F/412-591/Domain: serum albumin repeat homology <SA3>  
F/27-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,  
F/263/Binding site: bilirubin (Lys) #status predicted

Query Match 77.2%; Score 44; DB 1; Length 607;  
Best Local Similarity 66.7%; Pred. No. 2.8;  
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDLSKLE 12  
:|:|:|:|:|:|

Db 289 DQDLSKLE 300

RESULT 10

ABSHS  
serum albumin precursor - horse  
C/Species: Equus caballus (domestic horse)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C/Accession: S34053  
R/Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
Eur. J. Biochem. 215, 205-212, 1993  
A/Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A/Reference number: S34053; MUID:93345495; PMID:8344282  
A/Accession: S34053  
A/Molecule type: mRNA  
A/Residues: 1-607 <HOA>  
A/Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672  
A/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membr  
C/Suprafamily: serum albumin; serum albumin repeat homology  
C/Keywords: carrier protein; duplication; metal binding; plasma  
F/1-18/Domain: signal sequence #status predicted <SIG>  
F/19-24/Domain: propeptide #status predicted <PRO>  
F/25-607/Product: serum albumin #status predicted <MAT>  
F/220-393/Domain: serum albumin repeat homology <SA1>  
F/412-591/Domain: serum albumin repeat homology <SA2>  
F/27-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,  
F/263/Binding site: bilirubin (Lys) #status predicted

Query Match 75.4%; Score 43; DB 1; Length 607;  
Best Local Similarity 81.8%; Pred. No. 4.3;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ENQDSISSKXK 11
   :|||||
Db 289 ENQDSISSKXK 299

RESULT 11
JC5838
albumin - Mongolian jird
C:Species: Meriones unguiculatus (Mongolian jird)
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
C:Accession: JC5838
R:Yoshida, K.; Seto-Oshima, A.; Sinochara, H.
DNA Res. 4, 351-354, 1997
A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
A:Reference number: JC5838; MUID:9811663; PMID:9455485
A:Accession: JC5838
A:Molecule type: mRNA
A:Residues: 1-609 <YOS>
A:Cross-references: DBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2317278
A:Experimental source: liver
C:Superfamily: serum albumin; serum albumin repeat homology
F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 70.2%; Score 40; DB 2; Length 609;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKXK 11
   :|||||
Db 291 ENQDSISSKXK 301

RESULT 12
H36812
hypothetical protein ORF64 - saimirine herpesvirus 1 (strain 11)
C:Species: saimirine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C:Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 08-Oct-1999
C:Accession: H36812
R:Albrecht, J.
submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36806
A:Accession: H36812
A:Molecule type: DNA
A:Residues: 1-2469 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45687.1; PID:g60385
R:Albrecht, J. C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
J. Virol. 66, 5047-5059, 1992
A:Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688; PMID:1321287
A:Contents: annotation; protein-coding frames
A:Note: neither protein nor nucleotide sequence is given
C:Genetics:
A:Gene: 64

Query Match 70.2%; Score 40; DB 2; Length 2469;
Best Local Similarity 58.3%; Pred. No. 75;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKXKE 12
   :|||||
Db 373 EDDDNVTSKXKE 384

RESULT 13
A53195
afamin precursor - rat
N:Alternate names: alpha-albumin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-2000
C:Accession: A53195
R:Belanger, L.; Roy, S.; Allard, D.

```

```

J. Biol. Chem. 269, 5481-5484, 1994
A:Title: New albumin gene 3' adjacent to the alpha-1-fetoprotein locus.
A:Reference number: A53195; MUID:94164881; PMID:7509788
A:Accession: A53195
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-608 <BEL>
A:Cross-references: GB:X76456; NID:g456358; PIDN:CAA53394.1; PID:g456359
C:Genetics:
A:Introns: 30/1; 47/2; 91/3; 162/2
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-608/Product: afamin #status predicted <NAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-590/Domain: serum albumin repeat homology <SA3>

Query Match 68.4%; Score 39; DB 2; Length 608;
Best Local Similarity 88.9%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDSISSKXK 11
   :|||||
Db 292 QDSISSKXK 300

RESULT 14
T32033
hypothetical protein F41B5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32033
R:Dante, M.; Kramer, J.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F41B5.
A:Reference number: Z21115
A:Accession: T32033
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-238 <DAN>
A:Cross-references: EMBL:AF016676; PIDN:AAC25902.1; GSPDB:GN00023; CESP:F41B5.5
A:Experimental source: strain Bristol N2; clone F41B5
C:Genetics:
A:Gene: CESP:F41B5.5
A:Map position: 5
A:Introns: 77/1; 132/1

Query Match 66.7%; Score 38; DB 2; Length 238;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENQDSISSKXKE 12
   :|||||
Db 163 ENQDSAEAYLKE 174

RESULT 15
T34107
hypothetical protein C18C4.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34107
R:Gattung, S.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid C18C4.
A:Reference number: Z21478
A:Accession: T34107
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1091 <GAT>
A:Cross-references: EMBL:U55369; PIDN:AAC25825.1; GSPDB:GN00023; CESP:C18C4.5
A:Experimental source: strain Bristol N2; clone C18C4

```

C:Genetics:  
 A:Gene: CESP:C18C4.5  
 A:Map position: 5  
 A:Introns: 38/3; 95/3; 179/2; 259/3; 301/3; 419/2; 573/3; 613/1; 875/3; 920/2; 959/2; 98

Query Match 66.7%; Score 38; DB 2; Length 1091;  
 Best Local Similarity 58.3%; Pred. No. 74;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENQDSISSKLYE 12  
 :|||:|  
 Db 189 QNMDLSLSEKLS 200

Search completed: April 19, 2004, 12:02:28  
 Job time : 2.4626 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.875346 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_266\_277

Perfect score: 57  
Sequence: 1 ENQDSISSKLKE 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 57    | 100.0       | 600    | 1  | ALBU_MACMU  |
| 2          | 57    | 100.0       | 609    | 1  | ALBU_HUMAN  |
| 3          | 54    | 94.7        | 608    | 1  | ALBU_CANFA  |
| 4          | 54    | 94.7        | 608    | 1  | ALBU_FELCA  |
| 5          | 51    | 89.5        | 605    | 1  | ALBU_PIG    |
| 6          | 51    | 89.5        | 607    | 1  | ALBU_BOVIN  |
| 7          | 45    | 78.9        | 422    | 1  | MANA_HUMAN  |
| 8          | 44    | 77.2        | 599    | 1  | AFAM_HUMAN  |
| 9          | 44    | 77.2        | 607    | 1  | ALBU_SHEEP  |
| 10         | 43    | 75.4        | 607    | 1  | ALBU_HORSE  |
| 11         | 40    | 70.2        | 609    | 1  | ALBU_MERUN  |
| 12         | 40    | 70.2        | 2469   | 1  | TEGU_HSVSA  |
| 13         | 39    | 68.4        | 608    | 1  | AFAM_RAT    |
| 14         | 39    | 68.4        | 608    | 1  | ALBU_RABIT  |
| 15         | 39    | 68.4        | 611    | 1  | AFAM_MOUSE  |
| 16         | 37    | 64.9        | 416    | 1  | YTNL_BACSU  |
| 17         | 37    | 64.9        | 608    | 1  | ALBU_MOUSE  |
| 18         | 37    | 64.9        | 608    | 1  | ALBU_RAT    |
| 19         | 36    | 63.2        | 176    | 1  | GREP_CAMEL  |
| 20         | 36    | 63.2        | 1011   | 1  | SECA_PEA    |
| 21         | 36    | 63.2        | 1170   | 1  | YKDB_YEAST  |
| 22         | 36    | 63.2        | 1679   | 1  | GC2B_MOUSE  |
| 23         | 35    | 61.4        | 248    | 1  | PCRB_METTH  |
| 24         | 35    | 61.4        | 296    | 1  | YC32_METUA  |
| 25         | 35    | 61.4        | 416    | 1  | FDT_MOUSE   |
| 26         | 35    | 61.4        | 417    | 1  | FDT_HUMAN   |
| 27         | 35    | 61.4        | 504    | 1  | SIX1_YEAST  |
| 28         | 35    | 61.4        | 626    | 1  | CC23_YEAST  |
| 29         | 35    | 61.4        | 635    | 1  | DNAX_BORBU  |
| 30         | 35    | 61.4        | 637    | 1  | SYT_GLOB    |
| 31         | 35    | 61.4        | 676    | 1  | SPL1_COTUA  |
| 32         | 35    | 61.4        | 852    | 1  | RA50_THEME  |
| 33         | 35    | 61.4        | 1312   | 1  | L159_CAEEL  |

34 35 61.4 1324 1 SMC4\_SCHPO  
35 35 61.4 3680 1 DMD\_CANFA  
36 35 61.4 3685 1 DMD\_HUMAN  
37 34 59.6 183 1 ANF2\_XENLA  
38 34 59.6 187 1 ANF1\_XENLA  
39 34 59.6 217 1 EXPI\_ERWCA  
40 34 59.6 267 1 RECK\_STAEP  
41 34 59.6 354 1 HOM1\_HUMAN  
42 34 59.6 366 1 HOM1\_MOUSE  
43 34 59.6 366 1 HOM1\_RAT  
44 34 59.6 403 1 YME8\_YEAST  
45 34 59.6 439 1 ENGA\_THEME

#### ALIGNMENTS

RESULT 1  
ALBU\_MACMU STANDARD; PRT; 600 AA.  
AC Q28522;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor (Fragment).  
GN ALB.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93211971; PubMed=8460152;  
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,  
RA Dwulet J., Putnam F.W.;  
RT "cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding";  
RL Proc Natl Acad Sci U S A. 90:2409-2413(1993).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMIL; M90463; AAA36906.1; -.  
CC PIR; A47391; A47391.  
CC HSPP; P02768; IE7B.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport\_p1ot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
CC NON TER 1 1  
CC SIGNAL <1 10 BY SIMILARITY.  
CC PROPEP 11 16 BY SIMILARITY.  
CC CHAIN 17 600 SERUM ALBUMIN.  
CC DOMAIN 17 197 ALBUMIN 1.  
CC DOMAIN 204 389 ALBUMIN 2.  
CC DOMAIN 396 587 ALBUMIN 3.  
CC FT

P41004 schizosacch  
O97592 canis fam11  
P11532 homo sapien  
Q91617 xenopus lae  
Q81898 xenopus lae  
P33882 erwinia car  
Q8crv9 staphylococ  
Q86ym7 homo sapien  
Q92zy3 mus musculu  
Q92z14 rattus norv  
Q04697 saccharomyc  
Q9x1f8 thermotoga

FT METAL 19 19 COPPER (BY SIMILARITY).  
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).  
 FT DISULFID 69 78 BY SIMILARITY.  
 FT DISULFID 91 107 BY SIMILARITY.  
 FT DISULFID 106 117 BY SIMILARITY.  
 FT DISULFID 140 185 BY SIMILARITY.  
 FT DISULFID 184 193 BY SIMILARITY.  
 FT DISULFID 216 262 BY SIMILARITY.  
 FT DISULFID 261 269 BY SIMILARITY.  
 FT DISULFID 281 295 BY SIMILARITY.  
 FT DISULFID 294 305 BY SIMILARITY.  
 FT DISULFID 332 377 BY SIMILARITY.  
 FT DISULFID 376 385 BY SIMILARITY.  
 FT DISULFID 408 454 BY SIMILARITY.  
 FT DISULFID 453 464 BY SIMILARITY.  
 FT DISULFID 477 493 BY SIMILARITY.  
 FT DISULFID 492 503 BY SIMILARITY.  
 FT DISULFID 530 575 BY SIMILARITY.  
 FT DISULFID 574 583 BY SIMILARITY.  
 SQ SEQUENCE 600 AA; 67880 MW; B45C871A670E740B CRC64;  
 Query Match 100.0%; Score 57; DB 1; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 0.0078;  
 Matches 1; Conservative 0; Mismatches 0; Gaps 0;  
 Indels 0;  
 Oy 1 ENODSISSKLE 12  
 Db 282 ENODSISSKLE 293  
 RESULT 2  
 ALBU\_HUMAN  
 ID ALBU\_HUMAN STANDARD; PRT; 609 AA.  
 AC P02768; Q95574; Q13140; Q9P117; Q9UHS3; Q9UJZ0;  
 DT 21-JUL-1996 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RA MEDLINE=86196112; PubMed=3009475;  
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
 RA Beattie W.G., Dugaiczky A.,  
 RA "Molecular structure of the human albumin gene is revealed by  
 RT nucleotide sequence within q11-22 of chromosome 4.";  
 RL J. Biol. Chem. 261:6747-6757(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.; AND VARIANT LYS-420.  
 RX MEDLINE=82081882; PubMed=6171778;  
 RX Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.,  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 coli.";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A.; AND VARIANT GLY-121.  
 RX MEDLINE=82105994; PubMed=6275391;  
 RX Dugaiczky A., Law S.W., Dennison O.E.;  
 RA "Nucleotide sequence and the encoded amino acids of human serum  
 albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 TISSUE=Liver;  
 RC Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).

RC TISSUE=Petal liver;  
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
 Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
 RT "Functional prediction of the coding sequences of 121 new genes  
 deduced by analysis of cDNA clones from human fetal liver.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.; AND VARIANT HIROSHIMA-1 LYS-378.  
 RA Huang M.C., Wu H.T.;  
 RT "The cDNA sequences of human serum albumin.";  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 TISSUE=Liver, and Skeletal muscle;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner K.H., Schenck C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Malek J.A., Gunaratne P.H.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 25-609.  
 RX MEDLINE=76197907; PubMed=1225573;  
 RX Maloun B., Moravsek L., Kostka V.;  
 RT "Complete amino acid sequence of human serum albumin.";  
 RL FEBS Lett. 58:134-137(1975).  
 RN [9]  
 RP SEQUENCE OF 25-609.  
 RA Brown J.R., Shockley P., Behrens P.Q.;  
 RL (In) Bing D.H. (eds.);  
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
 RL Pergamon Press, New York (1979).  
 RN [10]  
 RP SEQUENCE OF 1-455 FROM N.A.  
 RC TISSUE=Liver;  
 RA Menaya J., Parrilla R., Ayuso M.S.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [11]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=86140099; PubMed=2419329;  
 RX Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
 RT "The human albumin gene. Characterization of the 5' and 3' flanking  
 regions and the polymorphic gene transcripts.";  
 RL J. Biol. Chem. 261:3244-3251(1986).  
 RN [12]  
 RP SEQUENCE OF 222-229.  
 RX MEDLINE=76257808; PubMed=955075;  
 RX Walker J.E.;  
 RT "Lysine residue 199 of human serum albumin is modified by  
 acetylsalicylic acid.";  
 RL FEBS Lett. 66:173-175(1976).  
 RN [13]  
 RP SEQUENCE OF 25-44 AND 480-499.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RX Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 1994.";

Electrophoresis 15:1459-1465 (1994).

[14] DISULFIDE BONDS.

RN SABER M.A., Stockbauer P., Moravsek L., Meloun B.;  
 RP MEDLINE=91316157; PubMed=1859851;  
 RA Peach R.J., Brennan S.O.;  
 RT "Disulfide bonds in human serum albumin.";  
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).

[15] BILIRUBIN-BINDING SITE.

RN JACOBSEN C.;  
 RP MEDLINE=78186630; PubMed=656055;  
 RA "Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin.";  
 RL Biochem. J. 171:453-459(1978).

[16] VARIANTS NAG-2 AND NAG-3.

RN MEDLINE=87157744; PubMed=3828358;  
 RA Brennan S.O., Herbert P.;  
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second domain of serum albumin.";  
 RL Biochim. Biophys. Acta 912:191-197(1987).

[17] VARIANTS NAG-2 AND NAG-3.

RN MEDLINE=89068523; PubMed=3479777;  
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M., Satoh C., Neel J.V.;  
 RT "Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).

[18] VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.

RN MEDLINE=89345611; PubMed=2762316;  
 RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M., Neel J.V., Sakurabayashi I., Putnam F.W.;  
 RT "Point substitutions in Japanese allcoalbumins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).

[19] VARIANTS MANUS; OSAKA; NAGOVY; FUKUOKA; HONOLULU AND NEW-GUINEA.

RN MEDLINE=90115905; PubMed=2404284;  
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
 RT "Point substitutions in albumin genetic variants from Asia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).

[20] DESCRIPTION OF VARIANT REDHILL.

RN MEDLINE=90115852; PubMed=2104980;  
 RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of human serum albumin whose precursor has an aberrant signal peptidase cleavage site.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).

[21] VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.

RN MEDLINE=91062352; PubMed=2247440;  
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J., Watkins S., Putnam F.W.;  
 RT "Mutations in genetic variants of human serum albumin found in Italy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).

[22] VARIANT VENEZIA.

RN MEDLINE=91296740; PubMed=2068071;  
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M., Minchiotti L., Putnam F.W.;  
 RT "A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).

[23] VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23; KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.

RN MEDLINE=92052189; PubMed=1946412;  
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E., Matsuda Y.-I., Anaki I., Putnam F.W.;  
 RT "Genetic variants of serum albumin in Americans and Japanese.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).

[24] VARIANT CASEBROOK ASN-518.

RN MEDLINE=91316157; PubMed=1859851;  
 RA Peach R.J., Brennan S.O.;  
 RT "Structural characterization of a glycoprotein variant of human serum albumin: albumin Casebrook (494 Asp-->Asn).";  
 RL Biochim. Biophys. Acta 1097:49-54(1991).

[25] VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.

RN MEDLINE=92190239; PubMed=1347703;  
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H., Rochu D., Porta F.;  
 RT "Two allcoalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 57; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 0.0079;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ENQDSISSKKE 12  
 DB 290 ENQDSISSKKE 301

RESULT 3

ALBU CANFA STANDARD; PRT; 608 AA.

ID AC P49822; Q77705; Q9TS24;  
 AC 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serum albumin precursor (Allergen Can f 3).  
 GN ALB.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;

[1] SEQUENCE FROM N.A.  
 RC STRAIN=Beagle; TISSUE=Liver;  
 RA Hilger C.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

[2] SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=20148667; PubMed=10669848;  
 RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H., Valenta R., Spitzauer S.;  
 RT "Escherichia coli expression and purification of recombinant dog albumin, a cross-reactive animal allergen.";  
 RL J. Allergy Clin. Immunol. 105:279-285(2000).

[3] SEQUENCE OF 25-48.  
 RX MEDLINE=75011422; PubMed=4414013;  
 RA Dixon J.W., Sarkar B.;  
 RT "Isolation, amino acid sequence and copper(II)-binding properties of peptide (1-24) of dog serum albumin.";  
 RL J. Biol. Chem. 249:5872-5877(1974).

[4] SEQUENCE OF 25-38.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).

[5] SEQUENCE OF 215-478 FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=94201492; PubMed=7512102;  
 RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P., Kuehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
 RT "Molecular characterization of dog albumin as a cross-reactive allergen.";

```

RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ133489; CAB64867.1; -.
DR EMBL; Y17737; CAB76841.1; -.
DR EMBL; S72946; AAB30434.1; -.
DR HSP; P02768; IE7B.
DR HSC-2DPAGE; P49822; DOG.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT CHAIN 19 24 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CONFLICT 1 26 MKWTFISLFLFSSAYSGLVRRRA -> MDT (IN
FT REF. 2).
FT CONFLICT 146 146 A -> R (IN REF. 2).
FT CONFLICT 206 206 I -> T (IN REF. 2).
FT CONFLICT 349 349 V -> A (IN REF. 2).
FT CONFLICT 359 359 S -> A (IN REF. 1).
FT CONFLICT 448 448 V -> VV (IN REF. 5).
FT CONFLICT 474 474 D -> E (IN REF. 1).
FT CONFLICT 474 474 D -> E (IN REF. 1).
FT SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;
Query Match 94.7%; Score 54; DB 1; Length 608;
Best local similarity 91.7%; Pred. No. 0.028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ENQDSISLKLKE 12
DB 290 ENQDSISLKLKE 301

```

```

RESULT 4
ALBU_FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN AUB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA MDLINE=96194924; PubMed=8647469;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin." ;
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT CHAIN 19 24 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT SEQUENCE 608 AA; 68659 MW; 07E629CAC5P60E5F CRC64;
SQ

```



```

FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 38556BD1A1F4FF CRC64;

Query Match 89.5%; Score 51; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 0.1;
Matches 10; Conservative 2; Mismatches 0; Indels

QY 1 ENQDSISKLKE 12
DB 287 ENQDTISTKLKE 298

RESULT 6
ALBU BOVIN STANDARD; PRT; 607 AA.
ID ALBU BOVIN
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelele
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; B
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID:9913;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. J.
RA Submitted (AUG-1991) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Barry T., Power S., Gannon F.;
RA TISSUE=Liver;
RA Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges
RA Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RA "The complete cDNA sequence of bovine serum albumin.";
RA Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGillivray R.A., Chung D.W., Davie E.W.;
RA "Biosynthesis of bovine plasma proteins in a cell-free sy
RT terminal sequence of preproalbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RA "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RA Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.

```

RX MEDLINE=82023364; PubMed=7283978;  
RA Reed R.G., Putnam F.W., Peters T. Jr.;  
RT "Sequence of residues 400-403 of bovine serum albumin.";  
RL Biochem. J. 191:867-868(1980).  
[9]  
RP SEQUENCE OF 19-28.  
RX MEDLINE=77134075; PubMed=843354;  
RA Patterson J.E., Geller D.M.;  
RT "Bovine microsomal albumin: amino terminal sequence of bovine  
proalbumin.";  
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
[10]  
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
RX MEDLINE=91083649; PubMed=2260975;  
RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;  
RT "Rapid confirmation and revision of the primary structure of bovine  
serum albumin by ESMS and revision of the Frit-FAB LC/MS.";  
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).  
[11]  
RP SEQUENCE OF 25-41.  
RX MEDLINE=88267456; PubMed=3389500;  
RA Hsieh J.C., Lin F.P., Tam M.F.;  
RT "Electroblotting onto glass-fiber filter from an analytical  
isoelectrofocusing gel: a preparative method for isolating proteins  
for N-terminal microsequencing.";  
RL Anal. Biochem. 170:1-8(1988).  
[12]  
RP SEQUENCE OF 437-451.  
RA Vilbois F.;  
RL Submitted (AUG-1998) to Swiss-Prot.  
[13]  
RP DISULFIDE BONDS.  
RA Brown J.R.;  
RL "Structure of serum albumin: disulfide bridges.";  
RL Fed. Proc. 33:1389-1389(1974).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M73993; AAA51411.1; -;  
CC EMBL; X58989; CAA41735.1; -;  
CC EMBL; Y17769; CAA76847.1; -;  
CC EMBL; AF542068; AAN17824.1; -;  
CC HSP; P02768; 1E7B.  
CC InterPro: IPR00264; Serum\_albumin.  
CC Pfam: PF00273; transport\_prot; 3.  
CC PRINTS; P00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
KW Polymorphism.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607 SERUM ALBUMIN.  
FT DOMAIN 25 204 ALBUMIN 1.  
FT DOMAIN 211 396 ALBUMIN 2.  
FT DOMAIN 403 594 ALBUMIN 3.  
FT METAL 27 27 COPPER (BY SIMILARITY).

FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
FT VARIANT 214 214  
FT CONFLICT 302 302  
FT CONFLICT 304 305  
FT CONFLICT 324 324  
FT CONFLICT 394 395  
FT CONFLICT 437 437  
FT CONFLICT 493 494  
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
A -> T. (IN REF. 6).  
C -> K (IN REF. 6).  
KP -> PC (IN REF. 6).  
N -> D (IN REF. 6).  
ST -> TS (IN REF. 6).  
K -> R (IN REF. 12).  
SE -> ES (IN REF. 6).  
Query Match 89.5%; Score 51; DB 1; Length 607;  
Best Local Similarity 83.3%; Pred. No. 0.1;  
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ENQDSISSKLE 12  
Db 289 DNQDTSSKLE 300  
RESULT 7  
MANA HUMAN  
ID MANA HUMAN STANDARD; PRT; 422 AA.  
AC P34949; 1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Mannose-6-phosphate isomerase (EC 5.3.1.8) (Phosphomannose isomerase)  
DE (PMI) (Phosphohexomutase).  
GN MPI OR PMI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Placenta, and Testis;  
RX MEDLINE=94139717; PubMed=8307007;  
RA Proudfoot A.E.I., Turcatti G., Wells T.N.C., Payton M.A., Smith D.J.;  
RT "Purification, cDNA cloning and heterologous expression of human  
phosphomannose isomerase.";  
RL Eur. J. Biochem. 219:415-423(1994).  
[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=20438347; PubMed=10980531;  
RA Scholten E., Dorland L., de Koning T.J., Van Diggelen O.P., Patterson M.,  
RA Huijman J.G.M., Marquardt T., Babovic-Vuksanovic D., Patterson M.,  
RA Intiaz F., Winchester B., Adamowicz M., Pronicka E., Freeze H.,  
RA Matthijs G.;  
RT "Genomic organization of the human phosphomannose isomerase (MPI) gene  
and mutation analysis in patients with congenital disorders of  
glycosylation type Ib (CDG-Ib).";  
RL Hum. Mutat. 16:247-252(2000).  
[3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Blood;  
RC MEDLINE=22388257; PubMed=12477932;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
RA Diatchenko L., Narusins K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
RN [4]  
RP VARIANTS CDG-IB LEU-101 AND THR-137.  
RX MEDLINE=98254476; PubMed=9585601;  
RA Jaeken J., Matthijs G., Saudubray J.-M., Dionisi-Vici C., Bertini E.,  
RA de Lonlay P., Henri H., Caruchon H., Schollen E., Van Schaftingen E.,  
RT "Phosphomannose isomerase deficiency: a carbohydrate-deficient  
RT glycoprotein syndrome with hepatic-intestinal presentation.";  
RL Am. J. Hum. Genet. 62:1535-1539 (1998).  
RN [5]  
RP VARIANT CDG-IB GLN-218.  
RX MEDLINE=98192611; PubMed=9525984;  
RA Niehus R., Hasilik A., Alton G., Koerner C., Schliebe-Sukumar M.,  
RA Koch H.G., Zimmer K.-P., Wu R., Harms E., Reiter K., von Figura K.,  
RA Freeze H.H., Harms H.K., Marquardt T.,  
RT "Carbohydrate-deficient glycoprotein syndrome type Ib: phosphomannose  
RT isomerase deficiency and mannose therapy.";  
RL J. Clin. Invest. 101:1414-1420 (1998).  
RN [6]  
RP VARIANTS CDG-IB THR-139 AND GLN-218.  
RX MEDLINE=21249093; PubMed=11350186;  
RA Westphal V., Kjaergaard S., Davis J.A., Peterson S.M., Skovby F.,  
RA Freeze H.H.;  
RT "Genetic and metabolic analysis of the first adult with congenital  
RT disorder of glycosylation type Ib: long-term outcome and effects of  
RT mannose supplementation.";  
RL Mol. Genet. Metab. 73:77-85 (2001).  
CC -!- FUNCTION: Involved in the synthesis of the GDP-mannose and  
CC dolichol-phosphate-mannose required for a number of critical  
CC mannose transfer reactions.  
CC -!- CATALYTIC ACTIVITY: D-mannose 6-phosphate = D-fructose 6-  
CC phosphate.  
CC -!- COFACTOR: Zinc (By similarity).  
CC -!- PATHWAY: Glycosylation; early steps of mannosylation.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- TISSUE SPECIFICITY: Expressed in all tissues, but more abundant in  
CC heart, brain and skeletal muscle.  
CC -!- DISEASE: Defects in MPI are the cause of congenital disorder of  
CC glycosylation type Ib (CDG-Ib) (MIM:602579); also known as  
CC carbohydrate-deficient glycoprotein syndrome type IB (CDGS-IB).  
CC CDG-IB is clinically characterized by protein-losing enteropathy,  
CC a gastrointestinal disorder. Biochemically it is characterized by  
CC hypoglycosylation of serum glycoproteins.  
CC -!- SIMILARITY: Belongs to the mannose-6-phosphate isomerase family 1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; X76057; CAA53657.1; -.  
DR EMBL; AF227218; AAF37697.1; -.

DR EMBL; AF227216; AAF37697.1; JOINED.  
DR EMBL; AF227217; AAF37697.1; JOINED.  
DR EMBL; BC046357; AAH46357.1; -.  
DR F1R; S41122; S41122.  
DR HSSP; P34948; LPMI.  
DR Genew; HGNC:7216; MPI.  
DR MIM; 154550; -.  
DR MIM; 602579; -.  
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; TAS.  
DR InterPro; IPR001250; Man6p\_isomerase.  
DR Pfam; PF01238; PMI type1; I.  
DR PRINTS; PR00714; MAN6PISMRASE.  
DR ProDom; PD004391; Man6p\_isomerase; 1.  
DR TIGRFAMs; TIGR00218; man6a; 1.  
DR PROSITE; PS00965; PMI\_1; 1.  
DR PROSITE; PS00966; PMI\_1.2; 1.  
KW Isomerase; Zinc; Disease mutation.  
FT INIT MET 0 PROBABLE.  
FT METAL 109 109 ZINC (BY SIMILARITY).  
FT METAL 111 111 ZINC (BY SIMILARITY).  
FT METAL 136 136 ZINC (BY SIMILARITY).  
FT METAL 275 275 ZINC (BY SIMILARITY).  
FT VARIANT 101 101 S -> L (in CDG-IB).  
FT VARIANT 137 137 /FTID=VAR\_012338.  
FT VARIANT 139 139 M -> T (in CDG-IB).  
FT VARIANT 139 139 I -> T (in CDG-IB).  
FT VARIANT 218 218 /FTID=VAR\_012345.  
FT VARIANT 218 218 R -> Q (in CDG-IB).  
FT SEQUENCE 422 AA; 46524 NM; /A450ABBF53722605 CRC64;  
SQ  
Query Match 78.9%; Score 45; DB 1; Length 422;  
Best Local Similarity 66.7%; Pred. No. 0.88;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ENQDSISSKLE 12  
Db 76 ENQDSLGSVKVD 87  
|||||:|:|:  
|:|:|:|:  
RESULT 8  
AFAM\_HUMAN STANDARD; PRT; 599 AA.  
ID AFAM\_HUMAN  
AC P43652;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE AFamin precursor (Alpha-albumin) (Alpha-Alb).  
GN AFM OR ALBA OR ALE2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC Tissue=Liver;  
RX MEDLINE=94299534; PubMed=7517938;  
RA Lichenstein H.S., Lyons D.E., Wurfel M.M., Johnson D.A.,  
RA McGinley M.D., Leidl J.C., Trollinger D.B., Mayer J.P.,  
RA Wright S.D., Zukowski M.M.;  
RT "Afinin is a new member of the albumin, alpha-fetoprotein, and  
RT vitamin D-binding protein gene family.";  
RL J. Biol. Chem. 269:18149-18154 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96240683; PubMed=8648639;  
RA Nishio H., Heiskanen M., Palotie A., Belanger L., Dugaiczak A.;  
RT "Random arrangement of the human serum albumin multigene family in  
RT the sub-centromeric region of 4q: evolution and chromosomal direction  
RT of transcription.";  
RL J. Mol. Biol. 259:113-119 (1996).  
RN [3]

SEQUENCE OF 1-69; 105-207 AND 560-599 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95180738; PubMed=7875606;  
RA Allard D., Gilbert S., Lamontagne A., Hamel D., Belanger L.;  
RT "Identification of rat alpha-albumin and cDNA cloning of its human  
ortholog";  
RL Gene 153:287-288(1995).  
CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: N-glycosylated.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L32140; AAA21612.1; -  
CC EMBL; U51243; AAC50720.1; -  
CC EMBL; L35486; AAA68197.1; -  
CC EMBL; L35497; AAA68198.1; -  
CC EMBL; L35498; AAA68199.1; -  
CC PIR; A54906; A54906.  
CC PIR; I39424; I39424.  
CC PIR; I39424; I39424.  
CC PIR; I39424; I39424.  
CC HSP; P02768; I39426.  
CC Gene; HGNC:316; AFM.  
CC MIM; 104145; -  
CC GO; GO:0005576; C:extracellular; TAS.  
CC InterPro; IPR00264; Serum albumin.  
CC Pfam; PF00273; transport\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Transport; Repeat; Glycoprotein; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 599 AFMIN.  
FT DOMAIN 22 205 ALBUMIN 1.  
FT DOMAIN 212 397 ALBUMIN 2.  
FT DOMAIN 404 593 ALBUMIN 3.  
FT DISULFID 77 86 BY SIMILARITY.  
FT DISULFID 99 114 BY SIMILARITY.  
FT DISULFID 113 124 BY SIMILARITY.  
FT DISULFID 148 193 BY SIMILARITY.  
FT DISULFID 182 201 BY SIMILARITY.  
FT DISULFID 224 270 BY SIMILARITY.  
FT DISULFID 269 277 BY SIMILARITY.  
FT DISULFID 289 303 BY SIMILARITY.  
FT DISULFID 302 313 BY SIMILARITY.  
FT DISULFID 340 385 BY SIMILARITY.  
FT DISULFID 384 393 BY SIMILARITY.  
FT DISULFID 416 462 BY SIMILARITY.  
FT DISULFID 461 470 BY SIMILARITY.  
FT DISULFID 483 499 BY SIMILARITY.  
FT DISULFID 498 509 BY SIMILARITY.  
FT DISULFID 536 581 BY SIMILARITY.  
FT DISULFID 580 589 BY SIMILARITY.  
FT CARBOHYD 33 33 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 599 AA; 69065 MW; D594E75E20D308AB CRC64;  
Query Match 77.2%; Score 44; DB 1; Length 599;  
Best Local Similarity 90.0%; Pred. No. 1.9;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 QDSISSKLKE 12

DB 292 QDSISSKLKE 301  
ALBU SHEEP STANDARD; PRT; 607 AA.  
RESULT 9  
ID ALBU SHEEP STANDARD; PRT; 607 AA.  
AC F14639; 201-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=90098888; PubMed=2602160;  
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;  
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin";  
RL Nucleic Acids Res. 17:10495-10495(1989).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X17055; CAA34903.1; -  
CC PIR; S06936; ABSHS.  
CC HSP; P02768; IE7B.  
CC InterPro; IPR00264; Serum albumin.  
CC Pfam; PF00273; transport\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 607 SERUM ALBUMIN.  
FT DOMAIN 25 204 ALBUMIN 1.  
FT DOMAIN 211 396 ALBUMIN 2.  
FT DOMAIN 403 594 ALBUMIN 3.  
FT METAL 27 27 COPPER (BY SIMILARITY).  
FT DISULFID 77 86 BY SIMILARITY.  
FT DISULFID 99 115 BY SIMILARITY.  
FT DISULFID 114 125 BY SIMILARITY.  
FT DISULFID 147 192 BY SIMILARITY.  
FT DISULFID 191 200 BY SIMILARITY.  
FT DISULFID 223 269 BY SIMILARITY.  
FT DISULFID 268 276 BY SIMILARITY.  
FT DISULFID 288 302 BY SIMILARITY.  
FT DISULFID 301 312 BY SIMILARITY.  
FT DISULFID 339 384 BY SIMILARITY.  
FT DISULFID 383 392 BY SIMILARITY.  
FT DISULFID 415 461 BY SIMILARITY.  
FT DISULFID 460 471 BY SIMILARITY.  
FT DISULFID 484 500 BY SIMILARITY.

```

FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 77.2%; Score 44; DB 1; Length 607;
Best Local Similarity 66.7%; Pred. No. 2;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLXE 12
: : : : :
Db 289 DHQDALSXLXE 300

RESULT 10
ALBU HORSE
ID ALBU HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]_TaxID=9796;
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
at 0.27-nm resolution."
RL Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X74045; CAA52194.1; -
PIR; S34053; ABHOS.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000284; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 BY SIMILARITY.
FT DOMAIN 25 204 SERUM ALBUMIN.
FT DOMAIN 211 396 ALBUMIN 1.
FT DOMAIN 403 594 ALBUMIN 2.
FT DOMAIN 27 27 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 85
FT DISULFID 99 115
FT DISULFID 114 125

```

```

FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 289 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 75.4%; Score 43; DB 1; Length 607;
Best Local Similarity 81.8%; Pred. No. 3;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLK 11
: : : : :
Db 289 EHQDSISGKLK 299

RESULT 11
ALBU MERUN
ID ALBU MERUN STANDARD; PRT; 609 AA.
AC Q35030;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]_TaxID=10047;
RP SEQUENCE FROM N.A.
RC STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=98116663; PubMed=9455485;
RA Yoshida K., Seto-Oshima A., Sinozawa H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
synthesis in the Mongolian gerbil, Meriones unguiculatus."
RL DNA Res. 4:351-354(1997).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AB006197; BAA21765.1; -
PIR; JC5838; JCS838.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.

```

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y64346; CAA45687.1; -  
DR EMBL; M86409; AAA46140.1; -  
DR InterPro; IPR006928; Herpes\_teg\_N.  
DR Pfam; PF04843; Herpes\_teg\_N.1.  
SQ SEQUENCE 2469 AA; 280165 MW; D2B4B8DC08644CDB CRC64;  
  
Query Match 70.2%; Score 40; DB 1; Length 2469;  
Best Local Similarity 58.3%; Pred. No. 45;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKLKE 12  
Db 373 EDDDNVTSKLKE 384  
|:|:|:|:|:|:|  
|:|:|:|:|:|:|  
  
RESULT 13  
AFAM\_RAT STANDARD; PRT; 608 AA.  
ID AFAM\_RAT  
AC P36953;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Afamin precursor (Alpha-albumin) (Alpha-Alb).  
GN AFM.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=94164881; PubMed=7509788;  
RA Belanger L., Roy S., Allard D.;  
RT "New albumin gene 3' adjacent to the alpha 1-fetoprotein locus";  
RL J. Biol. Chem. 269:5481-5484(1994).  
CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X76456; CAA53994.1; -  
CC PIR; A53195; A53195.  
CC HSPP; P02768; I37B.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC PRODOM; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 1.  
CC Transprot; Repeat; Glycoprotein; Signal.  
CC SIGNAL 1 21 BY SIMILARITY.  
CC CHAIN 1 21 APAMIN  
CC FT 22 608 ALBUMIN 1.  
CC FT DOMAIN 22 205 ALBUMIN 2.  
CC FT DOMAIN 212 397 ALBUMIN 3.  
CC FT DOMAIN 404 593 ALBUMIN 3.  
CC FT DISULFID 77 86 BY SIMILARITY.  
CC FT DISULFID 99 114 BY SIMILARITY.  
CC FT DISULFID 113 124 BY SIMILARITY.  
CC FT DISULFID 148 193 BY SIMILARITY.  
CC FT DISULFID 224 270 BY SIMILARITY.  
CC FT DISULFID 269 277 BY SIMILARITY.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 18 BY SIMILARITY.  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 609 SERUM ALBUMIN.  
FT DOMAIN 25 206 ALBUMIN 1.  
FT DOMAIN 213 398 ALBUMIN 2.  
FT DOMAIN 405 596 ALBUMIN 3.  
FT METAL 28 28 COPPER.  
FT DISULFID 78 87 BY SIMILARITY.  
FT DISULFID 100 116 BY SIMILARITY.  
FT DISULFID 115 126 BY SIMILARITY.  
FT DISULFID 149 194 BY SIMILARITY.  
FT DISULFID 193 202 BY SIMILARITY.  
FT DISULFID 225 271 BY SIMILARITY.  
FT DISULFID 270 278 BY SIMILARITY.  
FT DISULFID 290 304 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 341 386 BY SIMILARITY.  
FT DISULFID 385 394 BY SIMILARITY.  
FT DISULFID 417 463 BY SIMILARITY.  
FT DISULFID 462 473 BY SIMILARITY.  
FT DISULFID 486 502 BY SIMILARITY.  
FT DISULFID 501 512 BY SIMILARITY.  
FT DISULFID 539 584 BY SIMILARITY.  
FT DISULFID 583 592 BY SIMILARITY.  
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F57E1A48 CRC64;  
  
Query Match 70.2%; Score 40; DB 1; Length 609;  
Best Local Similarity 81.8%; Pred. No. 11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKLK 11  
Db 291 ENQASISSKLQ 301  
|:|:|:|:|:|:|  
|:|:|:|:|:|:|  
  
RESULT 12  
TEGU\_HSVSA STANDARD; PRT; 2469 AA.  
ID TEGU\_HSVSA  
AC Q01056;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE Probable large tegument protein.  
GN 64 OR ERF2.  
OS Herpesvirus saimiri (strain 11).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gamaherpesvirinae; Rhadinovirus.  
OX NCBI\_TaxID=10393;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92233688; PubMed=1321287;  
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,  
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,  
RA Honess R.W.;  
RT "Primary structure of the herpesvirus saimiri genome";  
RL J. Virol. 66:5047-5059(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92230228; PubMed=1314457;  
RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;  
RT "Analysis of nucleotide sequence of the rightmost 43 kbp of  
RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic  
RT organization between HVS and Epstein-Barr virus";  
RL Virology 188:296-310(1992).  
CC -!- FUNCTION: Tegument protein.  
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,  
CC BHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

|                       |  |           |           |                                     |  |
|-----------------------|--|-----------|-----------|-------------------------------------|--|
| FT                    | DISULFID   | 289       | 303       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 302       | 313       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 340       | 385       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 384       | 393       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 416       | 462       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 461       | 470       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 483       | 499       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 498       | 509       | BY SIMILARITY.                      |  |
| FT                    | DISULFID   | 580       | 589       | BY SIMILARITY.                      |  |
| FT                    | CARBOHYD   | 33        | 33        | N-LINKED (GLCNAC. . .) (POTENTIAL). |  |
| FT                    | CARBOHYD   | 109       | 109       | N-LINKED (GLCNAC. . .) (POTENTIAL). |  |
| FT                    | CARBOHYD   | 153       | 153       | N-LINKED (GLCNAC. . .) (POTENTIAL). |  |
| FT                    | CARBOHYD   | 402       | 402       | N-LINKED (GLCNAC. . .) (POTENTIAL). |  |
| FT                    | CARBOHYD   | 488       | 488       | N-LINKED (GLCNAC. . .) (POTENTIAL). |  |
| SQ                    | SEQUENCE   | 608 AA;   | 69335 MW; | F33151A6B8A07F6 CRC64;              |  |
| Query Match           |  |           |           |                                     | 68.4%; Score 39; DB 1; Length 608;                 |
| Best Local Similarity |  |           |           |                                     | 89.9%; Pred. No. 16;                               |
| Matches               |  |           |           |                                     | 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0; |
| Qy                    | 3 QDSISSKIK 11   |           |           |                                     |  |
| Db                    | 292 QDSISSKIK 300  |           |           |                                     |  |
| RESULT 14             |  |           |           |                                     |  |
| ALBU_RABIT            |  |           |           |                                     |  |
| ID                    | ALBU_RABIT   | STANDARD; | PRT;      | 608 AA.                             |  |
| AC                    | P49065;  |           |           |                                     |  |
| DT                    | 01-FEB-1996 (Rel. 33, Created)   |           |           |                                     |  |
| DT                    | 01-FEB-1996 (Rel. 33, Last sequence update)  |           |           |                                     |  |
| DT                    | 28-FEB-2003 (Rel. 41, Last annotation update)  |           |           |                                     |  |
| DE                    | Serum albumin precursor.   |           |           |                                     |  |
| GN                    | ALB.   |           |           |                                     |  |
| OS                    | Oryctolagus cuniculus (Rabbit).  |           |           |                                     |  |
| OC                    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |           |                                     |  |
| OC                    | Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  |           |           |                                     |  |
| OX                    | NCBI_TaxID=9986;   |           |           |                                     |  |
| RN                    | [1]  |           |           |                                     |  |
| RP                    | SEQUENCE FROM N.A.   |           |           |                                     |  |
| RC                    | SPRAIN-New Zealand white; TISSUE=Liver;  |           |           |                                     |  |
| RA                    | Shelfield W.P., Syed S., Schuyler P.D.;  |           |           |                                     |  |
| RL                    | Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  |           |           |                                     |  |
| CC                    | -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  |           |           |                                     |  |
| CC                    | binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  |           |           |                                     |  |
| CC                    | hormones, bilirubin and drugs. Its main function is the regulation   |           |           |                                     |  |
| CC                    | of the colloidal osmotic pressure of blood.  |           |           |                                     |  |
| CC                    | -!- SUBCELLULAR LOCATION: Secreted.  |           |           |                                     |  |
| CC                    | -!- TISSUE SPECIFICITY: Plasma.  |           |           |                                     |  |
| CC                    | -!- SIMILARITY: Belongs to the ALB/APP/VDB family.   |           |           |                                     |  |
| CC                    | -!- SIMILARITY: Contains 3 albumin domains.  |           |           |                                     |  |
| -----                 |  |           |           |                                     |  |
| CC                    | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |           |                                     |  |
| CC                    | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |           |                                     |  |
| CC                    | the European Bioinformatics Institute. There are no restrictions on its  |           |           |                                     |  |
| CC                    | use by non-profit institutions as long as its content is in no way   |           |           |                                     |  |
| CC                    | modified and this statement is not removed. Usage by and for commercial  |           |           |                                     |  |
| CC                    | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |           |                                     |  |
| CC                    | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |           |           |                                     |  |
| CC                    | -----  |           |           |                                     |  |
| DR                    | EMBL; U18344; AAB58347.1; .  |           |           |                                     |  |
| DR                    | HSSP; P02768; 1E7B.  |           |           |                                     |  |
| DR                    | InterPro; IPR000264; Serum_albumin.  |           |           |                                     |  |
| DR                    | Pfam; PF00273; transport_prot; 3.  |           |           |                                     |  |
| DR                    | PRINTS; PR00802; SERUMALBUMIN.   |           |           |                                     |  |
| DR                    | ProDom; PD002486; Serum_albumin; 1.  |           |           |                                     |  |
| DR                    | SMART; SM00103; ALBUMIN; 3.  |           |           |                                     |  |
| DR                    | PROSITE; PS00212; ALBUMIN; 3.  |           |           |                                     |  |
| KW                    | Metal-binding; Lipid-binding; Repeat; Signal; Copper.  |           |           |                                     |  |
| FT                    | SIGNAL   | 1         | 18        | BY SIMILARITY.                      |  |
| FT                    | PROPEP   | 19        | 24        | BY SIMILARITY.                      |  |
| FT                    | CHAIN  | 25        | 608       | SERUM ALBUMIN.                      |  |
| FT                    | DOMAIN   | 25        | 205       | ALBUMIN 1.                          |  |

|                       |  |           |           |                         |  |
|-----------------------|--|-----------|-----------|-------------------------|--|
| FT                    | DOMAIN   | 212       | 397       | ALBUMIN 2.              |  |
| FT                    | DOMAIN   | 404       | 595       | ALBUMIN 3.              |  |
| FT                    | METAL  | 27        | 27        | COPPER.                 |  |
| FT                    | DISULFID   | 77        | 86        | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 99        | 115       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 114       | 125       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 148       | 193       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 192       | 201       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 224       | 270       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 269       | 277       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 289       | 303       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 302       | 313       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 340       | 385       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 384       | 393       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 416       | 462       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 461       | 472       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 485       | 501       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 500       | 511       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 538       | 583       | BY SIMILARITY.          |  |
| FT                    | DISULFID   | 582       | 591       | BY SIMILARITY.          |  |
| SQ                    | SEQUENCE   | 608 AA;   | 68914 MW; | CF5E92647AAFE9A2 CRC64; |  |
| Query Match           |  |           |           |                         | 68.4%; Score 39; DB 1; Length 608;                 |
| Best Local Similarity |  |           |           |                         | 66.7%; Pred. No. 16;                               |
| Matches               |  |           |           |                         | 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0; |
| Qy                    | 1 ENQDSISSKKE 12   |           |           |                         |  |
| Db                    | 290 ENQETISSHLKE 301   |           |           |                         |  |
| RESULT 15             |  |           |           |                         |  |
| AFAM_MOUSE            |  |           |           |                         |  |
| ID                    | AFAM_MOUSE   | STANDARD; | PRT;      | 611 AA.                 |  |
| AC                    | O89020;  |           |           |                         |  |
| DT                    | 15-JUL-1999 (Rel. 38, Created)   |           |           |                         |  |
| DT                    | 15-JUL-1999 (Rel. 38, Last sequence update)  |           |           |                         |  |
| DT                    | 10-OCT-2003 (Rel. 42, Last annotation update)  |           |           |                         |  |
| DE                    | Afamin precursor (Alpha-albumin) (Alpha-Alb).  |           |           |                         |  |
| GN                    | AFM.   |           |           |                         |  |
| OS                    | Mus musculus (Mouse).  |           |           |                         |  |
| OC                    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |           |                         |  |
| OC                    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.   |           |           |                         |  |
| OX                    | NCBI_TaxID=10090;  |           |           |                         |  |
| RN                    | [1]  |           |           |                         |  |
| RP                    | SEQUENCE FROM N.A.   |           |           |                         |  |
| RC                    | TISSUE=Diaphragm;  |           |           |                         |  |
| RA                    | van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;   |           |           |                         |  |
| RL                    | Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  |           |           |                         |  |
| CC                    | -!- FUNCTION: Possible role in the transport of yet unknown ligand.  |           |           |                         |  |
| CC                    | -!- SUBCELLULAR LOCATION: Secreted.  |           |           |                         |  |
| CC                    | -!- SIMILARITY: Belongs to the ALB/APP/VDB family.   |           |           |                         |  |
| CC                    | -!- SIMILARITY: Contains 3 albumin domains.  |           |           |                         |  |
| -----                 |  |           |           |                         |  |
| CC                    | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |           |                         |  |
| CC                    | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |           |                         |  |
| CC                    | the European Bioinformatics Institute. There are no restrictions on its  |           |           |                         |  |
| CC                    | use by non-profit institutions as long as its content is in no way   |           |           |                         |  |
| CC                    | modified and this statement is not removed. Usage by and for commercial  |           |           |                         |  |
| CC                    | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |           |                         |  |
| CC                    | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |           |           |                         |  |
| CC                    | -----  |           |           |                         |  |
| DR                    | EMBL; AJ011080; CAA09471.1; .  |           |           |                         |  |
| DR                    | HSSP; P02768; 1E7B.  |           |           |                         |  |
| DR                    | InterPro; IPR000264; Serum_albumin.  |           |           |                         |  |
| DR                    | Pfam; PF00273; transport_prot; 3.  |           |           |                         |  |
| DR                    | PRINTS; PR00802; SERUMALBUMIN.   |           |           |                         |  |
| DR                    | ProDom; PD002486; Serum_albumin; 1.  |           |           |                         |  |
| DR                    | SMART; SM00103; ALBUMIN; 3.  |           |           |                         |  |
| DR                    | PROSITE; PS00212; ALBUMIN; 2.  |           |           |                         |  |
| KW                    | Transport; Repeat; Glycoprotein; Signal.   |           |           |                         |  |
| FT                    | SIGNAL   | 1         | 21        | BY SIMILARITY.          |  |

|             |         |           |                                     |
|-------------|---------|-----------|-------------------------------------|
| FT CHAIN    | 22      | 611       | AFAMIN.                             |
| FT DOMAIN   | 22      | 205       | ALBUMIN 1.                          |
| FT DOMAIN   | 212     | 397       | ALBUMIN 2.                          |
| FT DOMAIN   | 404     | 593       | ALBUMIN 3.                          |
| FT DISULFID | 77      | 86        | BY SIMILARITY.                      |
| FT DISULFID | 99      | 114       | BY SIMILARITY.                      |
| FT DISULFID | 113     | 124       | BY SIMILARITY.                      |
| FT DISULFID | 148     | 193       | BY SIMILARITY.                      |
| FT DISULFID | 224     | 270       | BY SIMILARITY.                      |
| FT DISULFID | 269     | 277       | BY SIMILARITY.                      |
| FT DISULFID | 289     | 303       | BY SIMILARITY.                      |
| FT DISULFID | 302     | 313       | BY SIMILARITY.                      |
| FT DISULFID | 340     | 385       | BY SIMILARITY.                      |
| FT DISULFID | 384     | 393       | BY SIMILARITY.                      |
| FT DISULFID | 416     | 462       | BY SIMILARITY.                      |
| FT DISULFID | 461     | 470       | BY SIMILARITY.                      |
| FT DISULFID | 483     | 499       | BY SIMILARITY.                      |
| FT DISULFID | 498     | 509       | BY SIMILARITY.                      |
| FT DISULFID | 580     | 589       | BY SIMILARITY.                      |
| FT CARBOHYD | 33      | 33        | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 109     | 109       | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 153     | 153       | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 402     | 402       | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 488     | 488       | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SQ SEQUENCE | 611 AA; | 69635 MW; | 39E46B6E723F89C8 CRC64;             |

Query Match 68.4%; Score 39; DB 1; Length 611;  
 Best Local Similarity 88.9%; Pred. No. 16;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

|    |                   |
|----|-------------------|
| QY | 3 QDSISSKLK 11    |
|    |                   |
| DB | 292 QDSISSKIK 300 |

Search completed: April 19, 2004, 11:52:51  
 Job time : 1.87535 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 4.63158 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_266\_277

Perfect score: 57  
Sequence: 1 ENQDSISSKLKE 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description        |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1          | 57    | 100.0       | 417    | 4  | Q86YGO | Q86YGO homo sapien |
| 2          | 54    | 94.7        | 584    | 6  | Q7YSG3 | Q7YSG3 felis silve |
| 3          | 45    | 78.9        | 267    | 4  | Q8NHZ6 | Q8NHZ6 homo sapien |
| 4          | 45    | 78.9        | 362    | 4  | Q96AB0 | Q96AB0 homo sapien |
| 5          | 45    | 78.9        | 423    | 6  | Q8HX22 | Q8HX22 macaca fasc |
| 6          | 41    | 71.9        | 74     | 2  | Q7WVJ9 | Q7WVJ9 uncultured  |
| 7          | 41    | 71.9        | 74     | 2  | Q7WVJ7 | Q7WVJ7 uncultured  |
| 8          | 41    | 71.9        | 74     | 2  | Q7WV16 | Q7WV16 uncultured  |
| 9          | 41    | 71.9        | 74     | 2  | Q7WVH6 | Q7WVH6 uncultured  |
| 10         | 41    | 71.9        | 74     | 2  | Q7WVH4 | Q7WVH4 uncultured  |
| 11         | 41    | 71.9        | 74     | 2  | Q7WVH3 | Q7WVH3 uncultured  |
| 12         | 41    | 71.9        | 74     | 2  | Q7WVG9 | Q7WVG9 uncultured  |
| 13         | 41    | 71.9        | 74     | 2  | Q7WRK7 | Q7WRK7 uncultured  |
| 14         | 41    | 71.9        | 74     | 2  | Q7WRK6 | Q7WRK6 uncultured  |
| 15         | 41    | 71.9        | 74     | 2  | Q7WRK5 | Q7WRK5 uncultured  |
| 16         | 40    | 70.2        | 74     | 2  | Q7WVK4 | Q7WVK4 uncultured  |

|    |    |      |      |           |                    |
|----|----|------|------|-----------|--------------------|
| 17 | 40 | 70.2 | 74   | 2 Q7WVK1  | Q7WVK1 uncultured  |
| 18 | 40 | 70.2 | 74   | 2 Q7WV08  | Q7WV08 uncultured  |
| 19 | 40 | 70.2 | 74   | 2 Q7WV06  | Q7WV06 uncultured  |
| 20 | 40 | 70.2 | 74   | 2 Q7WVJ5  | Q7WVJ5 uncultured  |
| 21 | 40 | 70.2 | 74   | 2 Q7WVJ4  | Q7WVJ4 uncultured  |
| 22 | 40 | 70.2 | 74   | 2 Q7WVJ3  | Q7WVJ3 uncultured  |
| 23 | 40 | 70.2 | 74   | 2 Q7WVJ1  | Q7WVJ1 uncultured  |
| 24 | 40 | 70.2 | 74   | 2 Q7WV17  | Q7WV17 uncultured  |
| 25 | 40 | 70.2 | 74   | 2 Q7WV15  | Q7WV15 uncultured  |
| 26 | 40 | 70.2 | 74   | 2 Q7WV13  | Q7WV13 uncultured  |
| 27 | 40 | 70.2 | 74   | 2 Q7WV12  | Q7WV12 uncultured  |
| 28 | 40 | 70.2 | 74   | 2 Q7WVH8  | Q7WVH8 uncultured  |
| 29 | 40 | 70.2 | 74   | 2 Q7WVH7  | Q7WVH7 uncultured  |
| 30 | 40 | 70.2 | 74   | 2 Q7WVH5  | Q7WVH5 uncultured  |
| 31 | 40 | 70.2 | 74   | 2 Q7WVH2  | Q7WVH2 uncultured  |
| 32 | 40 | 70.2 | 74   | 2 Q7WVH0  | Q7WVH0 uncultured  |
| 33 | 40 | 70.2 | 74   | 2 Q7WRK8  | Q7WRK8 uncultured  |
| 34 | 40 | 70.2 | 74   | 2 Q7WRK4  | Q7WRK4 uncultured  |
| 35 | 40 | 70.2 | 74   | 2 Q7WRK3  | Q7WRK3 uncultured  |
| 36 | 40 | 70.2 | 74   | 2 Q95VB7  | Q95VB7 schistosoma |
| 37 | 40 | 70.2 | 2477 | 12 Q80BL6 | Q80BL6 saimirine   |
| 38 | 39 | 68.4 | 74   | 2 Q7WVK8  | Q7WVK8 uncultured  |
| 39 | 39 | 68.4 | 74   | 2 Q7WVK5  | Q7WVK5 uncultured  |
| 40 | 39 | 68.4 | 74   | 2 Q7WVK3  | Q7WVK3 uncultured  |
| 41 | 39 | 68.4 | 74   | 2 Q7WVK2  | Q7WVK2 uncultured  |
| 42 | 39 | 68.4 | 74   | 2 Q7WVI9  | Q7WVI9 uncultured  |
| 43 | 39 | 68.4 | 74   | 2 Q7WVI8  | Q7WVI8 uncultured  |
| 44 | 39 | 68.4 | 74   | 2 Q7WVH9  | Q7WVH9 uncultured  |
| 45 | 39 | 68.4 | 74   | 2 Q7WVG6  | Q7WVG6 uncultured  |

ALIGNMENTS

RESULT 1  
Q86YGO PRELIMINARY; PRT; 417 AA.  
ID Q86YGO  
AC Q86YGO; (TREMELrel. 24, Created)  
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 25, Last annotation update)  
DE Similar to alpha-fetoprotein.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041789; AAH41789.1; -;  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR PRODOM; PD002486; Serum albumin; 1.  
DR SMART; SMC0103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E9D CRC64;

Query Match 100.0%; Score 57; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 0.056; 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0;  
Qy 1 ENQDSISSKLKE 12  
Db 98 ENQDSISSKLKE 109

RESULT 2

Q7YSG3 Q7YSG3 PRELIMINARY; PRT; 584 AA.  
AC Q7YSG3, 01-OCT-2003 (Tremblrel. 25, Created)  
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Albumin (Fragment).  
ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
[1]  
RN Similar to mannose-6-phosphate isomerase (EC 5.3.1.8) (Mannose-6-phosphate isomerase) (Phosphomannose isomerase) (PMI) (Phosphohexomutase).  
RP SEQUENCE FROM N.A.  
RC Tissue=Liver;  
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P., Rumpold H., Valenta R., Spitzauer S.,  
RT "Escherichia coli expression and purification of recombinant cat albumin: IGF recognition, induction of basophil activation and lymphoproliferative responses in atopic patients.",  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ487677; CAD32275.1; -.  
FT NON TER 1  
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;  
  
Query Match 94.7%; Score 54; DB 6; Length 584;  
Best Local Similarity 91.7%; Pred. No. 0.27; Mismatches 1; Indels 0; Gaps 0;  
Matches 11; Conservative 1;  
  
QY 1 ENQDSISXKKE 12  
|||||:||||:  
DB 266 ENQDSISXKKE 277  
  
RESULT 3  
Q8NHZ6 PRELIMINARY; PRT; 267 AA.  
AC Q8NHZ6; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Mannose phosphate isomerase isoform (EC 5.3.1.8) (Mannose-6-phosphate isomerase) (Phosphomannose isomerase) (PMI) (Phosphohexomutase).  
DE (Phosphohexomutase).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RC Tissue=Brain;  
RA Guo J.H., Yu L.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: D-MANNOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.  
CC -1- COFACTOR: ZINC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE MANNOSE-6-PHOSPHATE ISOMERASE FAMILY 1.  
DR EMBL; AF504648; AM28199.1; -.  
DR GO; GO:0016953; F:isomerase activity; IEA.  
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR001250; Man6p\_isomerase.  
DR Pfam; PF01238; PMI\_type1; 1.  
DR PRINTS; PR00714; MAN6PISMRASE.  
DR ProDom; PD004391; Man6p\_isomerase; 1.  
DR TIGRFAMs; TIGR00218; manA; 1.  
DR PROSITE; PS00965; PMI\_I\_1; 1.  
DR PROSITE; PS00966; PMI\_I\_2; 1.  
KW Isomerase; Zinc.  
SQ SEQUENCE 267 AA; 23699 MW; 322DCD06F3F92FE CRC64;  
  
Query Match 78.9%; Score 45; DB 4; Length 267;  
Best Local Similarity 66.7%; Pred. No. 5.3; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;  
  
QY 1 ENQDSISXKKE 12  
|||||:||||:  
DB 77 ENQDSISXKKE 88  
  
RESULT 5  
Q8HXX2 PRELIMINARY; PRT; 423 AA.  
AC Q8HXX2; 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Mannose-6-phosphate isomerase.  
GN MPI.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Macaca.  
OX NCBI\_TaxID=9541;  
[1]  
RN SEQUENCE FROM N.A.  
RC Tissue=Brain cerebellum cortex;  
RA Kusuda J., Osada N., Hida M., Sugano S., Hashimoto K.;  
RT "Isolation and characterization of cDNA for macaque neurological disease genes."  
RT

QY 1 ENQDSISXKKE 12  
|||||:||||:  
DB 57 ENQDSISXKKE 68  
  
RESULT 4  
Q96AB0 PRELIMINARY; PRT; 362 AA.  
AC Q96AB0; 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Similar to mannose phosphate isomerase (EC 5.3.1.8) (Mannose-6-phosphate isomerase) (Phosphomannose isomerase) (PMI) (Phosphohexomutase).  
DE (Phosphohexomutase).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RC Tissue=Astrocytoma;  
RA Strausberg R.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: D-MANNOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.  
CC -1- COFACTOR: ZINC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE MANNOSE-6-PHOSPHATE ISOMERASE FAMILY 1.  
DR EMBL; BC017351; AAH17351.1; Created.  
DR GO; GO:0016953; F:isomerase activity; IEA.  
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR001250; Man6p\_isomerase.  
DR Pfam; PF01238; PMI\_type1; 1.  
DR PRINTS; PR00714; MAN6PISMRASE.  
DR ProDom; PD004391; Man6p\_isomerase; 1.  
DR TIGRFAMs; TIGR00218; manA; 1.  
DR PROSITE; PS00965; PMI\_I\_1; 1.  
DR PROSITE; PS00966; PMI\_I\_2; 1.  
KW Isomerase; Zinc.  
SQ SEQUENCE 362 AA; 39834 MW; 58CA9B39BF20C459 CRC64;  
  
Query Match 78.9%; Score 45; DB 4; Length 362;  
Best Local Similarity 66.7%; Pred. No. 7.1; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;  
  
QY 1 ENQDSISXKKE 12  
|||||:||||:  
DB 77 ENQDSISXKKE 88  
  
RESULT 5  
Q8HXX2 PRELIMINARY; PRT; 423 AA.  
AC Q8HXX2; 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Mannose-6-phosphate isomerase.  
GN MPI.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Macaca.  
OX NCBI\_TaxID=9541;  
[1]  
RN SEQUENCE FROM N.A.  
RC Tissue=Brain cerebellum cortex;  
RA Kusuda J., Osada N., Hida M., Sugano S., Hashimoto K.;  
RT "Isolation and characterization of cDNA for macaque neurological disease genes."  
RT

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB083318; BAC20597.1; -  
DR GO: GO:0015853; P:isomerase activity; IEA.  
DR GO: GO:0004470; P:zinc ion binding; IEA.  
DR GO: GO:0008270; P:zinc ion binding; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPR001250; Man6p\_isomerase.  
DR Pfam: PF01238; P:MI type1; 1.  
DR PRINTS: PR00714; MANGPISNRASE.  
DR ProDom: PD004391; Man6p\_isomerase1; 1.  
DR TIGRFAMs: TIGR00218; man6a; 1.  
DR PROSITE: PS00965; P:MI\_1; 1.  
DR PROSITE: PS00966; P:MI\_1\_2; 1.  
KW isomerase.  
SQ SEQUENCE 423 AA; 46601 MW; 90330E53A0F9DBE CRC64;  
  
Query Match 78.9%; Score 45; DB 6; Length 423;  
Best Local Similarity 66.7%; Pred. No. 8.3;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKLE 12  
Db |||||: |||:  
77 ENQDSLGLSKVD 88  
  
RESULT 6  
ID Q7WVJ9 PRELIMINARY; PRT; 74 AA.  
AC Q7WVJ9;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CpeB (Fragment).  
GN CPEB.  
OS uncultured Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=159733;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22753450; PubMed=12871235;  
RA Steglich C., Post A.F., Hess W.R.;  
RT "Analysis of natural populations of Prochlorococcus spp. in the  
RL northern Red Sea using phycoerythrin gene sequences.";  
DR EMBL: AF438694; AAP97621.1; -  
FT NON\_TER 1  
FT NON\_TER 74  
SQ SEQUENCE 74 AA; 8305 MW; CFD43CFDEB8B5B94 CRC64;  
  
Query Match 71.9%; Score 41; DB 2; Length 74;  
Best Local Similarity 70.0%; Pred. No. 8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKL 10  
Db |||||: |||:  
3 ENQDSVNSKI 12  
  
RESULT 7  
ID Q7WVJ7 PRELIMINARY; PRT; 74 AA.  
AC Q7WVJ7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CpeB (Fragment).  
GN CPEB.  
OS uncultured Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=159733;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22753450; PubMed=12871235;  
RA Steglich C., Post A.F., Hess W.R.;  
RT "Analysis of natural populations of Prochlorococcus spp. in the  
RL northern Red Sea using phycoerythrin gene sequences.";  
DR EMBL: AF438694; AAP97621.1; -  
FT NON\_TER 1  
FT NON\_TER 74  
SQ SEQUENCE 74 AA; 8305 MW; CFD43CFDEB8B5B94 CRC64;  
  
Query Match 71.9%; Score 41; DB 2; Length 74;  
Best Local Similarity 70.0%; Pred. No. 8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKL 10  
Db |||||: |||:  
3 ENQDSVNSKI 12

RP SEQUENCE FROM N.A.  
RX MEDLINE=22753450; PubMed=12871235;  
RA Steglich C., Post A.F., Hess W.R.;  
RT "Analysis of natural populations of Prochlorococcus spp. in the  
RL northern Red Sea using phycoerythrin gene sequences.";  
DR EMBL: AF438701; AAP97628.1; -  
FT NON\_TER 1  
FT NON\_TER 74  
SQ SEQUENCE 74 AA; 8309 MW; 27A46DFDEB8D1989 CRC64;  
  
Query Match 71.9%; Score 41; DB 2; Length 74;  
Best Local Similarity 70.0%; Pred. No. 8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKL 10  
Db |||||: |||:  
3 ENQDSVNSKI 12  
  
RESULT 8  
ID Q7WV16 PRELIMINARY; PRT; 74 AA.  
AC Q7WV16;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CpeB (Fragment).  
GN CPEB.  
OS uncultured Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=159733;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22753450; PubMed=12871235;  
RA Steglich C., Post A.F., Hess W.R.;  
RT "Analysis of natural populations of Prochlorococcus spp. in the  
RL northern Red Sea using phycoerythrin gene sequences.";  
DR EMBL: AF438714; AAP97641.1; -  
FT NON\_TER 1  
FT NON\_TER 74  
SQ SEQUENCE 74 AA; 8252 MW; C8997CF7E18D139E CRC64;  
  
Query Match 71.9%; Score 41; DB 2; Length 74;  
Best Local Similarity 70.0%; Pred. No. 8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ENQDSISSKL 10  
Db |||||: |||:  
3 ENQDSVNSKI 12  
  
RESULT 9  
ID Q7WVH6 PRELIMINARY; PRT; 74 AA.  
AC Q7WVH6;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CpeB (Fragment).  
GN CPEB.  
OS uncultured Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=159733;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22753450; PubMed=12871235;  
RA Steglich C., Post A.F., Hess W.R.;  
RT "Analysis of natural populations of Prochlorococcus spp. in the  
RL northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).

DR EMBL; AF438734; AAP97661.1; -.

FT NON\_TER 1

FT NON\_TER 74 74

SQ SEQUENCE 74 AA; 8327 MW; D7B9B56DEB8E7994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;

Best Local Similarity 70.0%; Pred. No. 8;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10

Db 3 ENQDSVNSKI 12

RESULT 10

ID Q7WVH4

AC Q7WVH4 PRELIMINARY; PRT; 74 AA.

DT 01-OCT-2003

DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE CpeB (Fragment).

GN CPEB.

OS uncultured Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI\_TaxID=159733;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22753450; PubMed=12871235;

RA Steglich C., Post A.F., Hess W.R.;

RT "Analysis of natural populations of Prochlorococcus spp. in the

RT northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).

DR EMBL; AF438738; AAP97665.1; -.

FT NON\_TER 1

FT NON\_TER 74 74

SQ SEQUENCE 74 AA; 8332 MW; 24A77CFDEB9B02F9 CRC64;

Query Match

Best Local Similarity 71.9%; Score 41; DB 2; Length 74;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10

Db 3 ENQDSVNSKI 12

RESULT 11

ID Q7WVH3

AC Q7WVH3 PRELIMINARY; PRT; 74 AA.

DT 01-OCT-2003

DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE CpeB (Fragment).

GN CPEB.

OS uncultured Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI\_TaxID=159733;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22753450; PubMed=12871235;

RA Steglich C., Post A.F., Hess W.R.;

RT "Analysis of natural populations of Prochlorococcus spp. in the

RT northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).

DR EMBL; AF438739; AAP97666.1; -.

FT NON\_TER 1

FT NON\_TER 74 74

SQ SEQUENCE 74 AA; 8362 MW; 27A46DE040E0B2F9 CRC64;

Query Match

Best Local Similarity 71.9%; Score 41; DB 2; Length 74;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10

Db 3 ENQDSVNSKI 12

RESULT 12

ID Q7WVG9

AC Q7WVG9 PRELIMINARY; PRT; 74 AA.

DT 01-OCT-2003

DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE CpeB (Fragment).

GN CPEB.

OS uncultured Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI\_TaxID=159733;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22753450; PubMed=12871235;

RA Steglich C., Post A.F., Hess W.R.;

RT "Analysis of natural populations of Prochlorococcus spp. in the

RT northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).

DR EMBL; AF438745; AAP97672.1; -.

FT NON\_TER 1

FT NON\_TER 74 74

SQ SEQUENCE 74 AA; 8294 MW; D7A47CFDEB8E7994 CRC64;

Query Match

Best Local Similarity 71.9%; Score 41; DB 2; Length 74;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10

Db 3 ENQDSVNSKI 12

RESULT 13

ID Q7WRK7

AC Q7WRK7 PRELIMINARY; PRT; 74 AA.

DT 01-OCT-2003

DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE CpeB (Fragment).

GN CPEB.

OS uncultured Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI\_TaxID=159733;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22753450; PubMed=12871235;

RA Steglich C., Post A.F., Hess W.R.;

RT "Analysis of natural populations of Prochlorococcus spp. in the

RT northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).

DR EMBL; AF438700; AAP97627.1; -.

FT NON\_TER 1

FT NON\_TER 74 74

SQ SEQUENCE 74 AA; 8308 MW; D7A46DEB8E7994 CRC64;

Query Match

Best Local Similarity 71.9%; Score 41; DB 2; Length 74;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 ENQDSISKL 10  
 |||||:|:  
 3 ENQDSVNSKI 12

Search completed: April 19, 2004, 12:00:07  
 Job time : 4.63158 secs

QY 1 ENQDSISKL 10  
 |||||:|:  
 Db 3 ENQDSVNSKI 12

RESULT 14  
 Q7WRK6 PRELIMINARY; PRT; 74 AA.  
 AC Q7WRK6  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CpeB (Fragment).  
 GN CPEB.  
 OS uncultured Prochlorococcus sp.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCEI\_TaxID=159733;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22753450; PubMed=12871235;  
 RA Steglich C., Post A.F., Hess W.R.;  
 RT "Analysis of natural populations of Prochlorococcus spp. in the  
 northern Red Sea using phycoerythrin gene sequences.";  
 RL Environ. Microbiol. 5:681-690(2003).  
 DR EMBL; AF438683; AAP97610.1; -.  
 DR EMBL; AF438722; AAP97649.1; -.  
 DR EMBL; AF438728; AAP97655.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 74  
 SQ SEQUENCE 74 AA; 8356 MW; D7A4756DEB994994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;  
 Best Local Similarity 70.0%; Pred. No. 8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISKL 10  
 |||||:|:  
 Db 3 ENQDSVNSKI 12

RESULT 15  
 Q7WRK5 PRELIMINARY; PRT; 74 AA.  
 AC Q7WRK5  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CpeB (Fragment).  
 GN CPEB.  
 OS uncultured Prochlorococcus sp.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCEI\_TaxID=159733;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22753450; PubMed=12871235;  
 RA Steglich C., Post A.F., Hess W.R.;  
 RT "Analysis of natural populations of Prochlorococcus spp. in the  
 northern Red Sea using phycoerythrin gene sequences.";  
 RL Environ. Microbiol. 5:681-690(2003).  
 DR EMBL; AF438689; AAP97616.1; -.  
 DR EMBL; AF438736; AAP97663.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 74  
 SQ SEQUENCE 74 AA; 8322 MW; D7A47CFDEB994994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;  
 Best Local Similarity 70.0%; Pred. No. 8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISKL 10

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.07295 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_170\_176

Perfect score: 32

Sequence: 1 QAAKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCPUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 32    | 100.0       | 585    | 1  | US-08-153-799-14  |
| 2          | 32    | 100.0       | 585    | 1  | US-08-446-186A-3  |
| 3          | 32    | 100.0       | 585    | 2  | US-08-984-176-1   |
| 4          | 32    | 100.0       | 585    | 2  | US-08-702-572-2   |
| 5          | 32    | 100.0       | 585    | 3  | US-08-769-746-2   |
| 6          | 32    | 100.0       | 585    | 4  | US-10-153-064-5   |
| 7          | 32    | 100.0       | 609    | 1  | US-08-222-619-3   |
| 8          | 32    | 100.0       | 609    | 1  | US-08-433-037-4   |
| 9          | 32    | 100.0       | 609    | 4  | US-08-897-956A-2  |
| 10         | 32    | 100.0       | 609    | 4  | US-10-153-064-7   |
| 11         | 32    | 100.0       | 609    | 4  | US-09-976-594-977 |
| 12         | 32    | 100.0       | 609    | 5  | PCT-US95-04075-3  |
| 13         | 32    | 100.0       | 610    | 2  | US-08-797-689-2   |
| 14         | 32    | 100.0       | 610    | 4  | US-09-984-186-2   |
| 15         | 32    | 100.0       | 651    | 4  | US-10-153-064-133 |
| 16         | 32    | 100.0       | 652    | 4  | US-10-153-064-96  |
| 17         | 32    | 100.0       | 652    | 4  | US-10-153-064-99  |
| 18         | 32    | 100.0       | 652    | 4  | US-10-153-064-105 |
| 19         | 32    | 100.0       | 652    | 4  | US-10-153-064-132 |
| 20         | 32    | 100.0       | 653    | 4  | US-10-153-064-131 |
| 21         | 32    | 100.0       | 656    | 4  | US-10-153-064-130 |
| 22         | 32    | 100.0       | 660    | 4  | US-10-153-064-90  |
| 23         | 32    | 100.0       | 660    | 4  | US-10-153-064-93  |
| 24         | 32    | 100.0       | 668    | 4  | US-10-153-064-102 |
| 25         | 32    | 100.0       | 676    | 4  | US-10-153-064-95  |
| 26         | 32    | 100.0       | 676    | 4  | US-10-153-064-98  |
| 27         | 32    | 100.0       | 676    | 4  | US-10-153-064-104 |

Sequence 127, App  
Sequence 129, App  
Sequence 125, App  
Sequence 123, App  
Sequence 92, Appl  
Sequence 101, App  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 89, Appl  
Sequence 23494, A  
Sequence 81, Appl  
Sequence 81, Appl  
Sequence 81, Appl  
Sequence 81, Appl  
Sequence 81, Appl  
Sequence 81, Appl

28 32 100.0 676 4 US-10-153-064-127  
29 32 100.0 676 4 US-10-153-064-129  
30 32 100.0 677 4 US-10-153-064-125  
31 32 100.0 680 4 US-10-153-064-123  
32 32 100.0 684 4 US-10-153-064-92  
33 32 100.0 692 4 US-10-153-064-101  
34 32 100.0 783 1 US-08-256-938-2  
35 32 100.0 787 1 US-08-256-938-4  
36 32 100.0 787 2 US-08-797-689-16  
37 32 100.0 787 4 US-09-984-186-16  
38 32 100.0 978 4 US-08-897-956A-3  
39 32 100.0 1184 4 US-10-153-064-89  
40 29 93.6 565 4 US-09-252-931A-23494  
41 28 87.5 43 1 US-07-998-003A-81  
42 28 87.5 43 1 US-08-453-274B-81  
43 28 87.5 43 1 US-08-453-695A-81  
44 28 87.5 43 1 US-08-268-161A-81  
45 28 87.5 43 2 US-08-453-702A-81

#### ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No 576683  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 28-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 369..419
; OTHER INFORMATION: /note= "Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; OTHER INFORMATION: natural HSA"
US-08-153-799-14
Query Match 100.0%; Score 32; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 3
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1
Query Match 100.0%; Score 32; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 4
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:

```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2
Query Match 100.0%; Score 32; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2
Query Match 100.0%; Score 32; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match 100.0%; Score 32; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3
Query Match 100.0%; Score 32; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```



APPLICANT: Barr, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschopp, Uwe F.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DiGiglio, Frank S.  
REGISTRATION NUMBER: 31,345  
REFERENCE/DOCKET NUMBER: 9108Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 32; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 9  
US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 32; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 10  
US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PFS56  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 32; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 11  
US-09-976-594-977  
; Sequence 977, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 977  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1  
US-09-976-594-977

Query Match 100.0%; Score 32; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 12  
PCT-US95-04075-3  
; Sequence 3, Application PC/TUS9504075  
; GENERAL INFORMATION:  
; APPLICANT: AMGEN INC.

FILED DATE: 31-03-1992  
PRIOR APPLICATION DATA:

INFORMATION FOR SEQ

```

1  APPLICATION NUMBER: PCT/FR93/00085
2  FILING DATE: 28-JAN-1993
3  ATTORNEY/AGENT INFORMATION:
4  NAME: Smith Ph.D., Julie K.
5  REGISTRATION NUMBER: P-38,619
6  REFERENCE/DOCKET NUMBER: ST92006-US
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE: (610) 454-3839
9  TELEFAX: (610) 454-3808
10 INFORMATION FOR SEQ ID NO: 2:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 610 amino acids
13 TYPE: amino acid
14 TOPOLOGY: linear
15 MOLECULE TYPE: protein
16 US-08-797-689-2
17
18 Query Match 100.0%; Score 32; DB 2; Length 610;
19 Best Local Similarity 100.0%; Pred. No. 35;
20 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
21
22 Qy 1 QAAADKAA 7
23 |||||
24 Db 194 QAAADKAA 200
25
26 RESULT 14
27 US-09-984-186-2
28 ; Sequence 2, Application US/09984186
29 ; Patent No. 6686179
30 ; GENERAL INFORMATION:
31 ; APPLICANT: Fleer, Reinhard
32 ; Fournier, Alain
33 ; Guittion, Jean-Dominique
34 ; Jung, Gerard
35 ; Yeh, Patrice
36
37 TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
38 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
39 CONTAINING SAID POLYPEPTIDES
40
41 NUMBER OF SEQUENCES: 36
42 CORRESPONDENCE ADDRESS:
43 ADDRESSEE: Rhone-Poulenc Rorer Inc.
44 STREET: 500 Arcola Road, 3043
45 CITY: Collegeville
46 STATE: PA
47 COUNTRY: USA
48 ZIP: 19426
49
50 COMPUTER READABLE FORM:
51 MEDIUM TYPE: Floppy disk
52 COMPUTER: Macintosh
53 OPERATING SYSTEM: System 7.1
54 SOFTWARE: Word 5.1 (Patentin)
55
56 CURRENT APPLICATION DATA:
57 APPLICATION NUMBER: US/09/984,186
58 FILING DATE: 29-Oct-2001
59 CLASSIFICATION: <Unknown>
60
61 PRIOR APPLICATION DATA:
62 APPLICATION NUMBER: US/08/797,689
63 FILING DATE: 31-JAN-1997
64 APPLICATION NUMBER: US 08/256,927
65 FILING DATE: 28-JUL-1994
66 APPLICATION NUMBER: FR 92/01064
67 FILING DATE: 31-JAN-1992
68 APPLICATION NUMBER: PCT/FR93/00085
69 FILING DATE: 28-JAN-1993
70
71 ATTORNEY/AGENT INFORMATION:
72 NAME: Smith Ph.D., Julie K.
73 REGISTRATION NUMBER: P-38,619
74 REFERENCE/DOCKET NUMBER: ST92006-US
75 TELECOMMUNICATION INFORMATION:
76 TELEPHONE: (610) 454-3839
77 TELEFAX: (610) 454-3808
78
79 INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 100.0%; Score 32; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAADKAA 7  
Db 194 QAADKAA 200

RESULT 15  
US-10-153-064-133  
; Sequence 133, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 133  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-064-133

Query Match 100.0%; Score 32; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7  
Db 236 QAADKAA 242

Search completed: April 19, 2004, 12:05:19  
Job time : 1.07295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.731302 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLLLE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 33    | 100.0       | 250    | 2 F83609 | hypothetical prote |
| 2          | 33    | 100.0       | 265    | 2 I46986 | albumin - dog (fra |
| 3          | 33    | 100.0       | 379    | 2 A57477 | potassium channel  |
| 4          | 33    | 100.0       | 453    | 2 A05139 | serum albumin - mo |
| 5          | 33    | 100.0       | 600    | 2 A47391 | serum albumin prec |
| 6          | 33    | 100.0       | 605    | 1 ABPG3  | serum albumin prec |
| 7          | 33    | 100.0       | 607    | 1 ABBS   | serum albumin prec |
| 8          | 33    | 100.0       | 607    | 1 ABBS   | serum albumin prec |
| 9          | 33    | 100.0       | 607    | 1 ABBS   | serum albumin prec |
| 10         | 33    | 100.0       | 608    | 1 ABRTS  | serum albumin prec |
| 11         | 33    | 100.0       | 608    | 2 S57632 | serum albumin prec |
| 12         | 33    | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 13         | 33    | 100.0       | 609    | 2 JC5838 | albumin - Mongolia |
| 14         | 33    | 100.0       | 714    | 2 A53595 | ATP-dependent DNA  |
| 15         | 31    | 93.9        | 132    | 2 G72774 | hypothetical prote |
| 16         | 31    | 93.9        | 159    | 2 A83600 | phosphopantetheine |
| 17         | 31    | 93.9        | 334    | 2 D83750 | pyridoxal phosphat |
| 18         | 31    | 93.9        | 348    | 2 B75445 | conserved hypothet |
| 19         | 31    | 93.9        | 394    | 2 T32042 | hypothetical prote |
| 20         | 31    | 93.9        | 542    | 2 T08777 | probable protein k |
| 21         | 31    | 93.9        | 550    | 2 H71922 | ATP-dependent zinc |
| 22         | 31    | 93.9        | 559    | 2 A86604 | poly(3-hydroxyalka |
| 23         | 31    | 93.9        | 601    | 2 B81723 | conserved hypothet |
| 24         | 31    | 93.9        | 912    | 1 A53215 | protein kinase C ( |
| 25         | 31    | 93.9        | 918    | 1 I48719 | protein kinase C ( |
| 26         | 30    | 90.9        | 90     | 2 A05186 | hypothetical prote |
| 27         | 30    | 90.9        | 346    | 2 A95324 | hypothetical prote |
| 28         | 30    | 90.9        | 406    | 2 H84590 | hypothetical prote |
| 29         | 30    | 90.9        | 409    | 2 T25935 | hypothetical prote |

30 30 90.9 415 2 H84718  
31 30 90.9 435 2 F84083  
32 30 90.9 435 2 AC3237  
33 30 90.9 554 2 G75376  
34 30 90.9 560 2 C38604  
35 30 90.9 707 2 T02835  
36 30 90.9 720 2 T40900  
37 30 90.9 816 2 A49151  
38 30 90.9 1117 2 AE0075  
39 30 90.9 1149 2 T20891  
40 30 90.9 1163 2 S07137  
41 30 90.9 1271 2 T49009  
42 30 90.9 1289 1 RMXRR3  
43 30 90.9 1400 1 I38185  
44 30 90.9 1513 1 RNR2C2  
45 30 90.9 1527 1 RNZMB2

#### ALIGNMENTS

##### RESULT 1

F83609  
Hypothetical protein PA0278 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83609  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nucleotide 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F83609  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-250 <STO>  
A:Cross-references: GB:A004466; GB:A004091; NID:99946120; PIDN:AG03667.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0278

Query Match 100.0%; Score 33; DB 2; Length 250;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

DB 191 HGDLLLE 196

##### RESULT 2

I46986

albumin - dog (fragment)

C:Species: Canis lupus familiaris (dog)

C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 20-Aug-1999

C:Accession: I46986

R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner;

U:Allergy Clin. Immunol. 93, 614-627, 1994

A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.

A:Reference number: I46986; MUID:94201492; PMID:7512102

A:Accession: I46986

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-265 <SPI>

A:Cross-references: GB:S72946; NID:9633937; PIDN:AAB30434.1; PID:9633938

C:Superfamily: serum albumin; serum albumin repeat homology

F/7-180/Domain: serum albumin repeat homology <SA2>

Query Match 100.0%; Score 33; DB 2; Length 265;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 HGDLLLE 6
DB      57 HGDLLLE 62

RESULT 3
potassium channel K-AB-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Sep-1999
C:Accession: A57477
R:Takumi, T.; Ishii, T.; Horio, Y.; Morishige, K.I.; Takahashi, N.; Yamada, M.; Yamashita, J. Biol. Chem. 270, 16339-16346, 1995
A:Title: A novel ATP-dependent inward rectifier potassium channel expressed predominantly in the brain
A:Reference number: A57477; MUID:95332346; PMID:7608203
A:Accession: A57477
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-379 <TRK>
A:Cross-references: GB:X86818; NID:G939969; PIDN:CAA60501.1; PID:G939970
C:Superfamily: G protein-activated potassium channel protein
C:Keywords: ATP; transmembrane protein

Query Match      100.0%; Score 33; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      93 HGDLLLE 98

RESULT 4
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; I48638
R:Minighetti, P.P.; Law, S.W.; Dugaiczky, A. Mol. Biol. Evol. 2, 347-358, 1985
A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes
A:Reference number: A93055; MUID:89216123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <M>
A:Cross-references: GB:M6111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Bocaccio, C.; Deschatrette, J.; Meunier-Rotival, M. Gene 88, 181-186, 1990
A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the human genome
A:Reference number: I48638; MUID:90269606; PMID:1971802
A:Accession: I48638
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOCC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match      100.0%; Score 33; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      173 HGDLLLE 178

RESULT 5
A47391
serum albumin precursor - rhesus macaque

```

```

C:Species: Macaca mulatta (rhesus macaque)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F. Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A:Notes: sequence extracted from NCBI backbone (NCBIN:128280, NCBIPI:128281)
C:Superfamily: serum albumin; serum albumin repeat homology <SA1>
F:21-194/Domain: serum albumin repeat homology <SA2>
F:213-386/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match      100.0%; Score 33; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      263 HGDLLLE 268

RESULT 6
ABFGS
serum albumin precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Weinstock, J.; Baldwin, G.S. Nucleic Acids Res. 16, 9045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-605 <WEI>
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M. J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006
A:Molecule type: protein
A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>
A:Experimental source: dental enamel
A:Note: albumin and other serum proteins are also found in bone
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F:17-22/Domain: propeptide #status predicted <PRO>
F:23-605/Product: serum albumin #status predicted <WAT>
F:27-199/Domain: serum albumin repeat homology <SA1>
F:218-391/Domain: serum albumin repeat homology <SA2>
F:410-589/Domain: serum albumin repeat homology <SA3>
F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390, 4
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match      100.0%; Score 33; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      268 HGDLLLE 273

```

## RESULT 7

ABBS  
 serum albumin precursor [validated] - bovine  
 C;Alternate names: 67K protein; preproalbumin  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 18-Aug-2000  
 C;Accession: A38885; A36401; A91258; B60808; G10780; D45800; A26693; A90309; A91458; A94  
 R;Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.  
 submitted to the EMBL Data Library, August 1991  
 A;Description: Bovine serum albumin: cDNA sequence and expression.  
 A;Reference number: A38885  
 A;Accession: A38885  
 A;Molecule type: mRNA  
 A;Residues: 1-607 <HOL>  
 A;Cross-references: EMBL:M73215  
 R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
 Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
 A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumi  
 A;Reference number: A36401; MUID:1083649; PMID:2260975  
 A;Accession: A36401  
 A;Molecule type: protein  
 A;Residues: 25-41, 'H', '43-189', 'E', '191-213', 'T', '215-323', 'D', '325-393', 'TS', '396-607' <HIR>  
 R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
 Eur. J. Biochem. 98, 477-485, 1979  
 A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
 A;Reference number: A91258; MUID:80024278; PMID:488109  
 A;Accession: A91258  
 A;Molecule type: protein  
 A;Residues: 1-32 <MAG>  
 R;Haieh, J.C.; Lin, P.P.; Tam, M.F.  
 Anal. Biochem. 170, 1-8, 1988  
 A;Title: Electrophoretic transfer of an analytical isoelectrofocusing 9  
 A;Reference number: A60808; MUID:88267456; PMID:3389500  
 A;Accession: S10780  
 A;Molecule type: protein  
 A;Residues: 25-41 <HSI>  
 R;Strawich, E.; Glimcher, M.J.  
 Eur. J. Biochem. 191, 47-56, 1990  
 A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu  
 A;Reference number: S10780; MUID:90336641; PMID:2379503  
 A;Accession: S10780  
 A;Molecule type: protein  
 A;Residues: 25-41, 'H', '43-57', '59-64' <SWR>  
 R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
 J. Immunol. 143, 1680-1684, 1989  
 A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
 A;Reference number: A45800; MUID:89341406; PMID:2474609  
 A;Accession: D45800  
 A;Molecule type: protein  
 A;Residues: 163-172 <CAR>  
 R;Carraway, R.E.; Mitra, S.P.; Cochran, D.E.  
 J. Biol. Chem. 262, 5968-5973, 1987  
 A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe  
 A;Reference number: A26693; MUID:87194805; PMID:2437111  
 A;Accession: A26693  
 A;Molecule type: protein  
 A;Residues: 165-172, 'L' <CA2>  
 R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
 Biochem. J. 191, 867-868, 1980  
 A;Title: Sequence of residues 400-403 of bovine serum albumin.  
 A;Reference number: A90309; MUID:82023364; PMID:7283978  
 A;Accession: A90309  
 A;Molecule type: protein  
 A;Residues: 402-433 <REE>  
 R;Brown, J.R.  
 Fed. Proc. 34, 591, 1975  
 A;Title: Structure of bovine serum albumin.  
 A;Reference number: A91458  
 A;Accession: A91458  
 A;Molecule type: protein  
 A;Residues: 25-41, 'H', '43-117', 'EQ', '120-179', '181-189', 'E', '191-194', 'A', '196-213', 'T', '215-388', 'H

R;Brown, J.R.  
 submitted to the Atlas, April 1975  
 A;Reference number: A94551  
 A;Accession: A94551  
 A;Molecule type: Protein  
 A;Residues: 190-195 <BR2>  
 R;Brown, J.R.  
 Fed. Proc. 33, 1389, 1974  
 A;Reference number: A91457  
 A;Contents: annotation; disulfide bonds  
 R;Werlen, R.C.; Offord, R.E.; Rose, K.  
 Biochem. J. 302, 907-911, 1994  
 A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
 A;Reference number: S55232; MUID:9501935; PMID:7945219  
 A;Accession: S55232  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 529-536;569-572 <WER>  
 C;Superfamily: serum albumin; serum albumin repeat homology  
 C;Keywords: carrier protein; copper binding; duplication; Plasma  
 F;19-24/Domain: signal sequence #status experimental <SIG>  
 F;125-607/Product: serum albumin #status experimental <PRO>  
 F;220-393/Domain: serum albumin repeat homology <SA1>  
 F;220-393/Domain: serum albumin repeat homology <SA2>  
 F;412-591/Domain: serum albumin repeat homology <SA3>  
 F;27/Binding site: copper (His) #status predicted  
 F;77-86;99-115;114-125;147-192;191-200;223-269;268-276;288-302;301-312;339-384;383-392;4  
 Query Match 100.0%; Score 33; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HGDLLLE 6  
 DB 270 HGDLLLE 275  
 RESULT 8  
 ABBS  
 serum albumin precursor - sheep  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C;Accession: S06936  
 R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
 Nucleic Acids Res. 17, 10495, 1989  
 A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
 A;Reference number: S06936; MUID:90098888; PMID:2602160  
 A;Accession: S06936  
 A;Molecule type: mRNA  
 A;Residues: 1-607 <BRO>  
 A;Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387  
 C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
 teroid hormones (weak bonds with these hormones promote their transfer across the membra  
 C;Superfamily: carrier protein; duplication; metal binding; Plasma  
 C;Keywords: carrier protein; duplication; metal binding; Plasma  
 F;19-24/Domain: signal sequence #status predicted <SIG>  
 F;125-607/Product: serum albumin #status predicted <PRO>  
 F;220-393/Domain: serum albumin repeat homology <SA1>  
 F;220-393/Domain: serum albumin repeat homology <SA2>  
 F;412-591/Domain: serum albumin repeat homology <SA3>  
 F;27/Binding site: copper (His) #status predicted  
 F;77-86;99-115;114-125;147-192;191-200;223-269;268-276;288-302;301-312;339-384;383-392;4  
 F;263/Binding site: bilirubin (Lys) #status predicted  
 Query Match 100.0%; Score 33; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HGDLLLE 6  
 DB 270 HGDLLLE 275

## RESULT 9

ABHOS  
 serum albumin precursor - horse  
 C/Species: Equus caballus (domestic horse)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C/Accession: S34053  
 R/Ho, J.X.; Holowachuk, B.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
 Eur. J. Biochem. 215, 205-212, 1993  
 A/Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
 A/Reference number: S34053; MUID:93345495; PMID:8344282  
 A/Accession: S34053  
 A/Molecule type: mRNA  
 A/Residues: 1-607 <HOA>  
 A/Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672  
 C/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, ceroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
 C/Superfamily: serum albumin; serum albumin repeat homology  
 C/Keywords: carrier protein; duplication; metal binding; plasma  
 F/1-18/Domain: signal sequence #status predicted <SIG>  
 F/19-24/Domain: propeptide #status predicted <PRO>  
 F/25-607/Product: serum albumin #status predicted <MAT>  
 F/29-201/Domain: serum albumin repeat homology <SA1>  
 F/220-393/Domain: serum albumin repeat homology <SA2>  
 F/412-591/Domain: serum albumin repeat homology <SA3>  
 F/27/Binding site: copper (His) #status predicted  
 F/77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4  
 F/263/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 33; DB 1; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6

Db 270 HGDLLE 275

## RESULT 10

ABRTS  
 serum albumin precursor - rat  
 N/Alternate names: preproalbumin  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 31-May-1979 #sequence\_revision 31-May-1979 #text\_change 22-Jun-1999  
 C/Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233  
 R/Sargent, T.D.; Yang, M.; Bonner, J.  
 Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981  
 A/Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
 A/Reference number: A93872; MUID:81223722; PMID:7017712  
 A/Accession: A93872  
 A/Molecule type: mRNA  
 A/Residues: 1-608 <SAR>  
 A/Cross-references: GB:V01222; GB:J00698; NID:G55627; PIDN:CAA24532.1; PID:G55628  
 R/Stauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.  
 J. Biol. Chem. 252, 6846-6855, 1977  
 A/Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis  
 A/Reference number: A92211; MUID:77249657; PMID:893447  
 A/Note: Cleavages during protein maturation  
 A/Accession: A92211  
 A/Molecule type: protein  
 A/Residues: 1-38 <STR>  
 R/Isemura, S.; Ikenaka, T.  
 J. Biochem. 83, 35-48, 1978  
 A/Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage  
 A/Reference number: A91946; MUID:78109429; PMID:564345  
 A/Accession: A91946  
 A/Molecule type: protein  
 A/Residues: 25-222 <IS1>  
 R/Isemura, S.; Ikenaka, T.  
 J. Biochem. 79, 1183-1196, 1976  
 A/Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a  
 A/Reference number: A91940; MUID:76260153; PMID:956149

A/Accession: A91940  
 A/Molecule type: protein  
 A/Residues: 223-288;572-608 <IS2>  
 A/Note: 262-Leu was also found  
 R/Aovagi, Y.; Ikenaka, T.; Ichida, F.  
 Cancer Res. 38, 3483-3486, 1978  
 A/Title: Copper(II)-binding ability of human alpha-fetoprotein.  
 A/Reference number: A90758; MUID:79001617; PMID:80265  
 A/Contents: annotation; copper binding  
 R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
 J. Immunol. 143, 1680-1684, 1999  
 A/Title: Structures of histamine-releasing peptides formed by the action of acid protease  
 A/Reference number: A45800; MUID:89341406; PMID:2474609  
 A/Accession: C45800  
 A/Status: Preliminary  
 A/Molecule type: protein  
 A/Residues: 166-173 <CAR>  
 R/Heard, J.  
 Mol. Cell. Biol. 7, 2425-2434, 1987  
 A/Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved  
 A/Reference number: I57621; MUID:87286876; PMID:3475566  
 A/Accession: I57621  
 A/Status: Preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-5 <RES>  
 A/Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412  
 C/Superfamily: serum albumin; serum albumin repeat homology  
 C/Keywords: carrier protein; duplication; metal binding; plasma  
 F/1-18/Domain: signal sequence #status experimental <SIG>  
 F/19-24/Domain: propeptide #status experimental <PRO>  
 F/25-608/Product: serum albumin #status experimental <MAT>  
 F/29-201/Domain: serum albumin repeat homology <SA1>  
 F/221-394/Domain: serum albumin repeat homology <SA2>  
 F/413-592/Domain: serum albumin repeat homology <SA3>  
 F/27/Binding site: copper (His) #status experimental  
 F/77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

Query Match 100.0%; Score 33; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6

Db 271 HGDLLE 276

## RESULT 11

S57632  
 serum albumin precursor - cat  
 C/Species: Felis silvestris catus (domestic cat)  
 C/Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
 C/Accession: JC4660; S57632  
 R/Hilger, C.; Grigioni, F.; Hentges, F.  
 Gene 169, 295-296, 1996  
 A/Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
 A/Reference number: JC4660; MUID:96194824; PMID:8647469  
 A/Accession: JC4660  
 A/Molecule type: mRNA  
 A/Residues: 1-608 <H12>  
 A/Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
 A/Experimental source: liver  
 C/Comment: This protein is the major protein component in plasma. It functions as a mult  
 ein has 35 conserved cysteine residues.  
 C/Superfamily: serum albumin; serum albumin repeat homology  
 C/Keywords: liver; plasma  
 F/1-18/Domain: signal sequence #status predicted <SIG>  
 F/19-24/Domain: propeptide #status predicted <PRP>  
 F/25-608/Product: serum albumin #status predicted <MAT>  
 F/29-202/Domain: serum albumin repeat homology <SA1>  
 F/413-592/Domain: serum albumin repeat homology <SA2>  
 F/413-592/Domain: serum albumin repeat homology <SA3>

Query Match

100.0%; Score 33; DB 2; Length 608;

Best Local Similarity 100.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6  
Db 271 HGDLLE 276

RESULT 12

ABHUS  
serum albumin precursor [validated] - human  
N/Alternate names: preproalbumin  
N/Contains: kinetensin  
C/Species: Homo sapiens (man)  
C/Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000  
C/Accession: A93743; I59427; I59286; I59313; G01747; S55314; A91420; S06422; S36  
R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur  
Nucleic Acids Res. 9, 6103-6114, 1981  
A/Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli  
A/Reference number: A93743; MUID:82081882; PMID:6171778  
A/Accession: A93743  
A/Molecule type: mRNA  
A/Residues: 1-419, 'K', 421-609 <LAW>  
A/Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23  
R/Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A/Reference number: A93936; MUID:82105994; PMID:6275391  
A/Accession: A93936  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-609 <DUG>  
A/Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A/Reference number: I39427; MUID:86140099; PMID:2419329  
A/Accession: I39427  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-26 <URA>  
A/Cross-references: GB:M13075; NID:g178330; PIDN:AAA51688.1; PID:G553173  
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A/Reference number: I59286; MUID:94181575; PMID:8134387  
A/Accession: I59286  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 282-290, 'KSRFDLQ' <WAT>  
A/Cross-references: GB:S69192; NID:G546032; PIDN:AA830282.1; PID:G546033  
A/Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia  
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A/Reference number: I59313; MUID:94294404; PMID:8022807  
A/Accession: I59313  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 589-590, 'ALPRRVNLLIQVLP' <WAD>  
A/Cross-references: GB:S70799; NID:G547231; PIDN:AA831177.1; PID:G547232  
A/Note: This frame-shift variant is designated albumin Bazzano; four additional variants  
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A/Reference number: G08292  
A/Accession: G01747  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-455 <MEN>  
A/Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
A/Reference number: S55314; MUID:95275251; PMID:7755581

A/Accession: S55314  
A/Molecule type: protein  
A/Residues: 19-27 <LED>  
R/Meloun, B.; Moravek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A/Title: Complete amino acid sequence of human serum albumin.  
A/Reference number: A91420; MUID:76187907; PMID:1225573  
A/Accession: A91420  
A/Molecule type: protein  
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
R/Roehr, U.; Spittler, G.; Tripiel, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A/Reference number: S06422  
A/Note: this paper is in German, with an English abstract  
A/Accession: S06422  
A/Molecule type: protein  
A/Residues: 25-48 <ROB>  
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A/Title: Mass spectrometric identification of modifications to human serum albumin treat  
A/Reference number: S36882; MUID:93384321; PMID:8373198  
A/Accession: S36882  
A/Molecule type: protein  
A/Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <PIN>  
R/Kausler, E.; Spittler, G.  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A/Title: Chem. Hoppe-Seyler 372, 849-855, 1991  
A/Title: Bruchtaeucke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol.  
A/Reference number: S17599; MUID:92126241; PMID:1772598  
A/Accession: S17599  
A/Molecule type: protein  
A/Residues: 25-54, 354-357, 431-447 <KAU>  
A/Note: 49-Leu was also found  
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1694, 1989  
A/Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A/Reference number: A45800; MUID:89341406; PMID:2474609  
A/Accession: A45800  
A/Molecule type: protein  
A/Residues: 166-173 <CAR>  
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from peptin-tr  
A/Reference number: A03239; MUID:86242180; PMID:3087352  
A/Accession: A03239  
A/Molecule type: protein  
A/Residues: 166-173, 'L' <MOG>  
R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S  
Proc. Natl. Acad. Sci. U.S.A. 91, 8721-8725, 1990  
A/Title: Mutations in genetic variants of human serum albumin found in Italy.  
A/Reference number: A38255; MUID:91062352; PMID:2247440  
A/Accession: A38255  
A/Molecule type: protein  
A/Residues: 76-111 <GAL1>  
A/Accession: B38255  
A/Molecule type: protein  
A/Residues: 82-105, 'K', 107-110 <GAL2>  
A/Note: this variant is designated albumin Vibo Valentia  
A/Accession: A38255  
A/Molecule type: protein  
A/Residues: 76-83, 'K', 85-106 <GAL3>  
A/Note: this variant is designated albumin Torino  
R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tanni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A/Title: The structural characterization and bilirubin-binding properties of albumin Hei  
A/Reference number: S33298; MUID:93292504; PMID:8513793  
A/Accession: S33298  
A/Molecule type: protein  
A/Residues: 255-263, 'E', 265-281 <MIN1>  
A/Note: this variant is designated albumin Herborn  
R/Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A/Title: Two allalbumins with identical electrophoretic mobility are produced by diffe;



A:Reference number: S21078; MUID:92190239; PMID:1347703  
 A:Accession: S21078  
 A:Molecule type: protein  
 A:Residues: 354-356 'K' 358-378 <MIN2>  
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported, Nature 358, 209-215, 1992  
 R:He, X.M.; Carter, D.C.  
 A:Title: Atomic structure and chemistry of human serum albumin.  
 A:Reference number: A46756; MUID:92334427; PMID:1630489  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R:Brown, J.R.; Shockley, P.; Behrens, P.O.  
 A:Title: Disulfide bonds in human serum albumin.  
 A:Reference number: A90930  
 A:Contents: annotation; disulfide bonds  
 R:Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A:Reference number: A90299; MUID:78186630; PMID:656055  
 A:Contents: annotation; bilirubin-binding site  
 R:Peters, T.; Reed, R.G.  
 A:Title: Serum albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20, in Albumin: Structure, Biosynthesis, Function, and active sites.  
 A:Reference number: A94408  
 A:Contents: annotation; binding sites  
 R:Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A:Reference number: A90028; MUID:83279982; PMID:6192711  
 A:Contents: annotation; gene position  
 R:Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A:Reference number: A46755; MUID:76257808; PMID:955075  
 A:Contents: annotation  
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)  
 R:Bohnev, J.P.; Fonda, M.L.; Feldhoff, R.C.  
 FEBS Lett. 298, 246-248, 1992  
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate  
 A:Reference number: A56294; MUID:92183881; PMID:1544460  
 A:Contents: annotation  
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in P  
 atase activity  
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C:Genetics:  
 A:Gene: GDB:ALB  
 A:Cross-references: GDB:118990; OMIM:103600  
 A:Map position: 4q11-4q13  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 F:19-24/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-609/Product: serum albumin #status experimental <MPT>  
 F:29-202/Domain: serum albumin repeat homology <SA1>  
 F:166-174/Product: kinetensin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:27/Binding site: copper (His) #status predicted  
 F:77-86,99-115,114-125,148-193,192-201,224-269-277,289-303,302-313,340-385,384-393,4  
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Db 271 HGDLLE 276  
 RESULT 13  
 JC5838  
 albumin - Mongolian jird  
 C:Species: Meriones unguiculatus (Mongolian jird)  
 C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
 C:Accession: JC5838  
 R:Yoshida, K.; Seto-Ohashima, A.; Sinohara, H.  
 DNA Res. 4, 351-354, 1997  
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the  
 A:Reference number: JC5838; MUID:98116663; PMID:9455485  
 A:Accession: JC5838  
 A:Molecule type: mRNA  
 A:Residues: 1-609 <YOS>  
 A:Cross-references: DDBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278  
 A:Experimental source: liver  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:222-395/Domain: serum albumin repeat homology <SA2>  
 Query Match 100.0%; Score 33; DB 2; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HGDLLE 6  
 Db 272 HGDLLE 277  
 RESULT 14  
 AE3595  
 ATP-dependent DNA helicase recg (EC 3.6.1.-) [imported] - Brucella melitensis (strain 1  
 C:Species: Brucella melitensis  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
 C:Accession: AE3595  
 R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
 ; Marur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes,  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
 A:Reference number: AB3252; PMID:11756688  
 A:Accession: AE3595  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-714 <XUR>  
 A:Cross-references: GB:AE080918; PIDN:AAL53928.1; PID:gl7984872; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BME110686  
 A:Map position: 11  
 C:Superfamily: DNA helicase recG  
 C:Keywords: Hydrolase  
 Query Match 100.0%; Score 33; DB 2; Length 714;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HGDLLE 6  
 Db 666 HGDLLE 671  
 RESULT 15  
 G72774  
 hypothetical protein APE0185 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C:Accession: G72774  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
 DNA Res. 6, 93-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
 A:Reference number: A7450; MUID:99310339; PMID:10382966

A;Accession: G72774  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-132 <KAW>  
 A;Cross-references: DDBJ:AP000058; NID:gs103388; PIDN:BAA79097.1; PID:dl042873; PID:gs103388  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APE0185  
 C;Superfamily: Aeropyrum pernix hypothetical protein APE0185

Query Match 93.9%; Score 31; DB 2; Length 132;  
 Best Local Similarity 83.3%; Pred. NO. 18;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLE 6  
 Db 111 HGDLE 116

Search completed: April 19, 2004, 12:02:27  
 Job time : 2.7313 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.437673 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_247\_252

Perfect score: 33

Sequence: 1 HGDLL6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description         |
|------------|-------|-------------|--------|--------------|---------------------|
| 1          | 33    | 100.0       | 379    | 1 IRKA HUMAN | P78508 homo sapien  |
| 2          | 33    | 100.0       | 379    | 1 IRKA MOUSE | P49633 mus musculus |
| 3          | 33    | 100.0       | 379    | 1 IRKA RAT   | P49655 r atp-sensi  |
| 4          | 33    | 100.0       | 600    | 1 ALBU MACMU | Q28522 macaca mula  |
| 5          | 33    | 100.0       | 605    | 1 ALBU FIG   | P08835 sus scrofa   |
| 6          | 33    | 100.0       | 607    | 1 ALBU BOVIN | P02769 bos taurus   |
| 7          | 33    | 100.0       | 607    | 1 ALBU HORSE | P35747 equus caball |
| 8          | 33    | 100.0       | 607    | 1 ALBU SHEEP | P14639 ovis aries   |
| 9          | 33    | 100.0       | 608    | 1 ALBU CANFA | P49822 canis famil  |
| 10         | 33    | 100.0       | 608    | 1 ALBU FELCA | P49064 felis silve  |
| 11         | 33    | 100.0       | 608    | 1 ALBU MOUSE | P07724 mus musculus |
| 12         | 33    | 100.0       | 608    | 1 ALBU RABIT | P49065 oryctolagus  |
| 13         | 33    | 100.0       | 608    | 1 ALBU RAT   | P02770 rattus norv  |
| 14         | 33    | 100.0       | 609    | 1 ALBU HUMAN | P02768 homo sapien  |
| 15         | 33    | 100.0       | 609    | 1 ALBU MERUN | O35090 meriones un  |
| 16         | 33    | 100.0       | 941    | 1 CDAC HUMAN | Q9un75 homo sapien  |
| 17         | 33    | 100.0       | 948    | 1 CDAC HUMAN | Q9y5h9 homo sapien  |
| 18         | 33    | 100.0       | 948    | 1 CDAA HUMAN | Q9y5i2 homo sapien  |
| 19         | 33    | 100.0       | 949    | 1 CDAB HUMAN | Q9y5i1 homo sapien  |
| 20         | 33    | 100.0       | 950    | 1 CD33 HUMAN | Q9y5h8 homo sapien  |
| 21         | 33    | 100.0       | 950    | 1 CDAD HUMAN | Q9y5i0 homo sapien  |
| 22         | 31    | 93.9        | 159    | 1 COAD PSEAE | Q916d1 pseudomonas  |
| 23         | 31    | 93.9        | 280    | 1 XRC2 HUMAN | Q43543 homo sapien  |
| 24         | 31    | 93.9        | 314    | 1 PRNA CLOTE | Q892r2 clostridium  |
| 25         | 31    | 93.9        | 334    | 1 PDVA BACHD | Q9rc88 bacillus ha  |
| 26         | 31    | 93.9        | 393    | 1 SET8 HUMAN | Q9aqrl homo sapien  |
| 27         | 31    | 93.9        | 559    | 1 PHAA PSEOL | P28494 pseudomonas  |
| 28         | 31    | 93.9        | 601    | 1 Y248 CHLMU | Q9p160 chlamydia m  |
| 29         | 31    | 93.9        | 878    | 1 KPCC HUMAN | Q9b2l6 homo sapien  |
| 30         | 31    | 93.9        | 890    | 1 KPCN HUMAN | Q94806 homo sapien  |
| 31         | 31    | 93.9        | 912    | 1 KPCM HUMAN | Q15139 homo sapien  |
| 32         | 31    | 93.9        | 918    | 1 KPCM MOUSE | Q62101 mus musculus |
| 33         | 30    | 90.9        | 159    | 1 COAD_PSESM | Q88ah3 pseudomonas  |

#### ALIGNMENTS

##### RESULT 1

| ID | IRKA HUMAN   | STANDARD; | PRT; | 379 AA. |
|----|--|-----------|------|---------|
| AC | P78508; Q92808;  |           |      |         |
| DT | 01-NOV-1997 (Rel. 35, Created)   |           |      |         |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update)  |           |      |         |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update)  |           |      |         |
| DE | ATP-sensitive inward rectifier potassium channel 10 (Potassium   |           |      |         |
| DE | channel, inwardly rectifying, subfamily J, member 10) (Inward  |           |      |         |
| DE | rectifier K+ channel Kir1.2) (ATP-dependent inwardly rectifying  |           |      |         |
| DE | potassium channel Kir4.1).   |           |      |         |
| GN | KCNJ10.  |           |      |         |
| OS | Homo sapiens (Human).  |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |      |         |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.   |           |      |         |
| OX | NCBI_TaxID=9606;   |           |      |         |
| RN | [1]  |           |      |         |
| RC | SEQUENCE FROM N.A.   |           |      |         |
| RP | TISSUE=Kidney;   |           |      |         |
| RX | MEDLINE=97150765; PubMed=8995301;  |           |      |         |
| RA | Shuck M.E., Piser T.M., Bock J.H., Slightom J.L., Lee K.S.,  |           |      |         |
| RA | Bienkowski M.J.;   |           |      |         |
| RT | "Cloning and characterization of two K+ inward rectifier (Kir) 1.1   |           |      |         |
| RL | potassium channel homologs from human kidney (Kir1.2 and Kir1.3).";  |           |      |         |
| RL | J. Biol. Chem. 272:586-593(1997).  |           |      |         |
| RN | [2]  |           |      |         |
| RC | SEQUENCE FROM N.A.   |           |      |         |
| RP | TISSUE=Cerebellum;   |           |      |         |
| RA | Schoots O., van Tol H.H.M.;  |           |      |         |
| RL | Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.  |           |      |         |
| CC | -1- FUNCTION: MAY BE RESPONSIBLE FOR POTASSIUM BUFFERING ACTION OF   |           |      |         |
| CC | GLIAL CELLS IN THE BRAIN. INWARD RECTIFIER K(+) CHANNELS ARE   |           |      |         |
| CC | CHARACTERIZED BY A GREATER TENDANCY TO ALLOW POTASSIUM TO FLOW   |           |      |         |
| CC | INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEPENDANCE IS   |           |      |         |
| CC | REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS  |           |      |         |
| CC | EXTERNAL K(+) IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING  |           |      |         |
| CC | SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS  |           |      |         |
| CC | MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL  |           |      |         |
| CC | MAGNESIUM. CAN BE BLOCKED BY EXTRACELLULAR BA(2+) AND CS(+) (BY  |           |      |         |
| CC | SIMILARITY).   |           |      |         |
| CC | -1- SUBUNIT: SEEMS TO FORM HETERODIMER WITH KIR5.1/KCNJ16 (BY  |           |      |         |
| CC | SIMILARITY).   |           |      |         |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein.   |           |      |         |
| CC | -1- SIMILARITY: Belongs to the inward rectifier-type potassium channel   |           |      |         |
| CC | family.  |           |      |         |
| CC | -----  |           |      |         |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |      |         |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |      |         |
| CC | the European Bioinformatics Institute. There are no restrictions on its  |           |      |         |
| CC | use by non-profit institutions as long as its content is in no way   |           |      |         |
| CC | modified and this statement is not removed. Usage by and for commercial  |           |      |         |
| CC | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |      |         |
| CC | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |           |      |         |
| CC | -----  |           |      |         |
| DR | EMBL; U73192; AAC50923.1; -.   |           |      |         |
| DR | EMBL; U73193; AAC50924.1; -.   |           |      |         |

Q88cq7 pseudomonas  
P26496 pseudomonas  
P12227 plasm sativ  
P11079 reovirus (t  
P56764 arabidopsis  
Q8thv5 sinapis alb  
Q9mtm3 oenothera h  
P38550 nicotiana t  
Q04912 homo sapien  
P12093 oryza sativ  
P16025 zea mays (m  
Q88be2 pseudomonas

34 30 90.9 161 1 COAD\_PSEPK  
35 30 90.9 560 1 PHAC\_PSEOL  
36 30 90.9 1163 1 RP0D\_PEA  
37 30 90.9 1289 1 MCE\_REOVD  
38 30 90.9 1376 1 RP0D\_ARATH  
39 30 90.9 1384 1 RP0D\_SINAL  
40 30 90.9 1386 1 RP0D\_OENHO  
41 30 90.9 1388 1 RP0D\_TOBAC  
42 30 90.9 1400 1 RON\_HUMAN  
43 30 90.9 1513 1 RP0D\_ORYSA  
44 30 90.9 1527 1 RP0D\_MAIZE  
45 29 87.9 214 1 KGUA\_PSESM

```

DR EMBL; U52155; AAB07046.1; -.
DR Genew; HGNC:6256; KCNJ10.
DR MIM; 602208; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0015272; P:ATP-activated inward rectifier potassium ch. .; TAS.
DR GO; GO:0006813; P:potassium ion transport; TAS.
DR InterPro; IPR001838; K-channel_IR.
DR InterPro; IPR001622; K-channel_pore.
DR Pfam; PF01007; IRK; 1.
DR PRINTS; PR01320; KIRCHANNEL.
DR ProDom; PD001103; K-channel_IR; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW ATP-binding; Potassium transport.
FT DOMAIN 1 69 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 70 92 M1 (POTENTIAL).
FT DOMAIN 118 134 H5 (Pore-forming) (POTENTIAL).
FT TRANSMEM 143 167 M2 (POTENTIAL).
FT DOMAIN 168 379 CYTOPLASMIC (POTENTIAL).
FT SITE 158 158 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).
FT NP_BIND 210 217 ATP (POTENTIAL).
FT CONFLICT 50 50 L -> P (IN REF. 2).
FT CONFLICT 166 166 L -> Q (IN REF. 2).
FT CONFLICT 271 271 R -> C (IN REF. 2).
SQ SEQUENCE 379 AA; 42508 MW; D9DA013FF4003533 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLE 6
DB 93 HGDLLE 98

RESULT 2
IRKA_MOUSE STANDARD; PRT; 379 AA.
ID IRKA_MOUSE AC Q9JMG3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-sensitive inward rectifier potassium channel 10 (Potassium channel, inwardly rectifying, subfamily J, member 10) (Inward rectifier K+ channel Kir4.1).
DE Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP INANOB A., Takahashi K., Tanemoto M., Fujita A., Kurachi Y.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1, and ICR; TISSUE=Brain; MEDLINE=21100968; PubMed=11169792;
RX Li L., Head V., Timpe L.C.;
RT Identification of an inward rectifier potassium channel gene expressed in mouse cortical astrocytes.;
RL Glia 33:57-71(2001).
CC -1- FUNCTION: MAY BE RESPONSIBLE FOR POTASSIUM BUFFERING ACTION OF GLIAL CELLS IN THE BRAIN. INWARD RECTIFIER K(+) CHANNELS ARE CHARACTERIZED BY A GREATER TENDENCY TO ALLOW POTASSIUM TO FLOW INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEPENDANCE IS REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS EXTERNAL K(+) IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL MAGNESIUM. CAN BE BLOCKED BY EXTRACELLULAR BA(2+) AND CS(+) (BY SIMILARITY).
CC

```

---

```

CC -1- SUBUNIT: SEEMS TO FORM HETERODIMER WITH KIRS.1/KCNJ16 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the inward rectifier-type potassium channel family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AB039879; BAA92432.1; -.
CC EMBL; AF322631; AAG42845.1; -.
CC MGD; MGI:1194504; Kcnj10.
CC InterPro; IPR001838; K-channel_IR.
CC InterPro; IPR001622; K-channel_pore.
CC Pfam; PF01007; IRK; 1.
CC PRINTS; PR01320; KIRCHANNEL.
CC ProDom; PD001103; K-channel_IR; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW ATP-binding; Potassium transport.
FT DOMAIN 1 69 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 70 92 M1 (POTENTIAL).
FT DOMAIN 118 134 H5 (Pore-forming) (POTENTIAL).
FT TRANSMEM 143 167 M2 (POTENTIAL).
FT DOMAIN 168 379 CYTOPLASMIC (POTENTIAL).
FT SITE 158 158 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).
FT NP_BIND 210 217 ATP (POTENTIAL).
FT SEQUENCE 379 AA; 42432 MW; 7FF08446B7F43453 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLE 6
DB 93 HGDLLE 98

RESULT 3
IRKA_RAT STANDARD; PRT; 379 AA.
ID IRKA_RAT AC P49655; Q62790;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-sensitive inward rectifier potassium channel 10 (Potassium channel, inwardly rectifying, subfamily J, member 10) (ATP-sensitive inward rectifier potassium channel KIR4.1) (KIR4.1) (BIR10) (Brain-specific inwardly rectifying K+ channel 1) (BIRK1).
DE Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP STRAIN=Sprague-Dawley; TISSUE=Brain; MEDLINE=95332346; PubMed=7608203;
RX Takumi T., Ishii T., Horio Y., Morishige K.-I., Takahashi N., Yamada M., Yamashita T., Kiyama H., Sohmiya K., Nakanishi S., Kurachi Y.;
RT "A novel ATP-dependent inward rectifier potassium channel expressed predominantly in glial cells.";
RL J. Biol. Chem. 270:16339-16346(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;

```



```
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67980 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 33; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLLE 6
Db 263 HGDLLLE 268

RESULT 5
ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S.; Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X12422; CAA30970.1; --
CC EMBL; M35787; AAA30988.1; --
CC PIR; S01382; ABPGS.
CC HSP; P02768; IE7H.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
KW NON_PEP 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
```

```
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match 100.0%; Score 33; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLLE 6
Db 268 HGDLLLE 273

RESULT 6
ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W.; Stoltenberg J.K.; Reed R.G.; Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T.; Power S.; Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C.; Grigioni F.; de Beaufort C.; Michel G.; Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T.; Huang M.C.;
RL "The complete cDNA sequence of bovine serum albumin.";
RN Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGillivray R.T.A.; Chung D.W.; Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RL terminal sequence of prealbumin.";
RN Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
```

RA Brown J.R.;  
RT "Structure of bovine serum albumin.";  
RL Fed. Proc. 34:591-591(1975).  
RN [7]  
RP REVISIONS TO 190-195.  
RA Brown J.R.;  
RL Submitted (APR-1975) to the PIR data bank.  
RN [8]  
RP SEQUENCE OF 402-433.  
RX MEDLINE=82023364; PubMed=7283978;  
RA Reed R.G., Putnam F.W., Peters T. Jr.;  
RT "Sequence of residues 400-403 of bovine serum albumin.";  
RL Biochem. J. 191:867-868(1980).  
RN [9]  
RP SEQUENCE OF 19-28.  
RX MEDLINE=77134075; PubMed=843354;  
RA Patterson J.E., Geller D.M.;  
RT "Bovine microsomal albumin: amino terminal sequence of bovine  
proalbumin.";  
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
RN [10]  
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
RX MEDLINE=91083649; PubMed=2260375;  
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;  
RT "Rapid confirmation and revision of the primary structure of bovine  
serum albumin by ESIMS and Revision of Frit-FAB LC/MS.";  
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).  
RN [11]  
RP SEQUENCE OF 25-41.  
RX MEDLINE=88267456; PubMed=3389500;  
RA Hsieh J.C., Lin F.P., Lam M.F.;  
RT "Electroblotting onto glass-fiber filter from an analytical  
isoelectrofocusing gel: a preparative method for isolating proteins  
for N-terminal microsequencing.";  
RL Anal. Biochem. 170:1-8(1988).  
RN [12]  
RP SEQUENCE OF 437-451.  
RA Vilbois F.;  
RL Submitted (AUG-1998) to Swiss-Prot.  
RN [13]  
RP DISULFIDE BONDS.  
RA Brown J.R.;  
RL "Structure of serum albumin: disulfide bridges.";  
RL Fed. Proc. 33:1389-1389(1974).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M73993; AAA51411.1; -  
DR EMBL; X58989; CAA41735.1; -  
DR EMBL; Y17759; CAA76847.1; -  
DR EMBL; AF542068; AAN17824.1; -  
DR HSSP; P02768; IE7B.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PRC0802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN\_3.  
DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
KW Polymorphism. 1 18  
FT SIGNAL 19 24  
FT PROPEP 25 607  
FT CHAIN 25 204  
FT DOMAIN 211 396  
FT DOMAIN 403 594  
FT METAL 27 27  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
FT VARIANT 214 214  
FT CONFLICT 302 302  
FT CONFLICT 304 305  
FT CONFLICT 324 324  
FT CONFLICT 394 395  
FT CONFLICT 437 437  
FT CONFLICT 493 494  
FT SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
SQ  
  
Query Match 100.0%; Score 33; DB 1; Length 607;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative C; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HGDLLLE 6  
DB 270 HGDLLLE 275  
  
RESULT 7  
ALBU HORSE  
ID -ALBU\_HORSE STANDARD; PRT; 607 AA.  
AC P35747;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Equ c 3).  
GN ALB.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
RC TISSUE=Liver;  
RX MEDLINE=93345495; PubMed=8344282;  
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
at 0.27-nm resolution.";  
RL Eur. J. Biochem. 215:205-212(1993).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.

```

Brown W.M., Dieglelewska K.M., Foreman R.C., Saunders N.R.;
"Nucleotide and deduced amino acid sequence of sheep serum albumin.";
Nucleic Acids Res. 17:10495-10495 (1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC CC -!- SUBCELLULAR LOCATION: Secreted.
CC CC -!- TISSUE SPECIFICITY: Plasma.
CC CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC CC -!- SIMILARITY: Contains 3 albumin domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
ENBL; X17055; CAA34903.1; -.
DR DR PIR; S06936; ABSHS.
DR DR HSSP; P02768; 1E7B.
DR DR InterPro; IPR000264; Serum_albumin.
DR DR Pfam; PF00273; transport_prot; 3.
DR DR PRINTS; PR00802; SERUMALBUMIN.
DR DR ProDom; PD002486; Serum_albumin; 1.
DR DR SMART; SM00103; ALBUMIN; 3.
DR DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;
Query Match 100.0%; Score 33; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy 1 HGDLL 6
Db 270 HGDLL 275
-----
RESULT 9
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 20-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).

```



GN ALB.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Beagle; TISSUE=Liver;  
RC Hilger C.; (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RL Submitted (12-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=20148667; PubMed=10669848;  
RA Pandjaitan B., Swoboda I., Brandesky-Pichler P., Rumpold H.,  
RA Valenta R., Spitzauer S.;  
RT "Escherichia coli expression and purification of recombinant dog  
RT albumin, a cross-reactive animal allergen.";  
RL J. Allergy Clin. Immunol. 105:279-285(2000).  
RN [3]  
RN SEQUENCE OF 25-48.  
RX MEDLINE=75011422; PubMed=4414013;  
RA Dixon J.W., Sakkar B.;  
RT "Isolation, amino acid sequence and copper(II)-binding properties of  
RT peptide (1-24) of dog serum albumin.";  
RL J. Biol. Chem. 249:5872-5877(1974).  
RN [4]  
RN SEQUENCE OF 25-38.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RT dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
RN [5]  
RN SEQUENCE OF 215-478 FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=94201492; PubMed=7512102;  
RA Spitzauer S., Schweiger C., Speer W.R., Pandjaitan B., Valent P.,  
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
RT "Molecular characterization of dog albumin as a cross-reactive  
RT allergen.";  
RL J. Allergy Clin. Immunol. 93:614-627(1994).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ133489; CAB64867.1; -.  
CC EMBL; Y17737; CA476841.1; -.  
CC EMBL; S72946; AAB30434.1; -.  
CC HSP; P02768; 1E7B.  
CC HSC-2DPAGE; P49822; DOG.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transprot\_prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608  
FT DOMAIN 26 205  
FT DOMAIN 212 397  
FT DOMAIN 404 595  
FT METAL 27 27  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 340 385  
FT DISULFID 384 393  
FT DISULFID 416 462  
FT DISULFID 461 472  
FT DISULFID 485 501  
FT DISULFID 500 511  
FT DISULFID 538 583  
FT DISULFID 582 591  
FT CONFLICT 1 26  
FT CONFLICT 146 146  
FT CONFLICT 206 206  
FT CONFLICT 349 349  
FT CONFLICT 359 359  
FT CONFLICT 448 448  
FT CONFLICT 474 474  
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;  
  
Query Match 100.0%; Score 33; DB 1; Length 608;  
Best Local Similarity 100.0%; Pred. No. 16; Indels 0; Gaps 0;  
Matches 6; Conservative  
  
QY 1 HGDLLE 6  
DB 271 HGDLLE 276  
  
RESULT 10  
ALBU\_FELCA STANDARD; PRT; 608 AA.  
AC P49064;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Fel d 2).  
GN ALB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=96194824; PubMed=8647469;  
RA Hilger C., Grigioni F., Kohnen M., Hertges F.;  
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";  
RL Gene 163:295-296(1996).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

DR EMBL; X84842; CAA59279.1; -;  
 DR PIR; JC4660; S57632.  
 DR HSSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport Prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.

DR SMART; SW00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.

DR KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608  
 FT DOMAIN 25 205  
 FT DOMAIN 212 397  
 FT DOMAIN 404 595  
 FT METAL 27 27

FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 148 193  
 FT DISULFID 192 201  
 FT DISULFID 224 277  
 FT DISULFID 289 303  
 FT DISULFID 302 313  
 FT DISULFID 340 385  
 FT DISULFID 384 393  
 FT DISULFID 416 462  
 FT DISULFID 461 472  
 FT DISULFID 485 501  
 FT DISULFID 500 511  
 FT DISULFID 538 583  
 FT DISULFID 582 591

SQ SEQUENCE 608 AA; 68659 MW; 07B629CAC5F60E5F CRC64;  
 Query Match 100.0%; Score 33; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLE 6  
 |||||  
 Db 271 HGDLLE 276

RESULT 11  
 ALBU MOUSE STANDARD; PRT; 608 AA.  
 ID ALBU MOUSE  
 AC P07724; Q61802;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB OR ALB1 OR ALB-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RA van Reeth T., Cabant P., Dreze P., Szpirer J., Szpirer C.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Tongue;  
 CC

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Sono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Pletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Rutterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 99-515 FROM N.A.  
 RC MEDLINE=89216123; PubMed=2452956;  
 RA Minghetti P.P., Law S.W., Dugaiczky A.;  
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that  
 of pseudogenes.";  
 RL Mol. Biol. Evol. 2:347-358(1985).  
 RN [5]  
 RP SEQUENCE OF 477-551 FROM N.A.  
 RC STRAIN=BALB/c;  
 RC MEDLINE=90269606; PubMed=1971802;  
 RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;  
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat  
 located in the mouse serum albumin-encoding gene.";  
 RL Gene 88:181-186(1990).  
 RN [6]  
 RP SEQUENCE OF 25-44.  
 RC TISSUE=Liver;  
 RC MEDLINE=93162044; PubMed=1285668;  
 RA Giometti C.S., Taylor J., Tollaksen S.L.;  
 RT "Mouse liver protein database: a catalog of proteins detected by two-  
 dimensional gel electrophoresis.";  
 RL Electrophoresis 13:970-991(1992).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC

CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AJ011413; CAA09617.1; --  
 CC EMBL; AK010025; BAB26650.1; --  
 CC EMBL; BC049971; AAH49971.1; --  
 CC EMBL; M16111; AAA37190.1; --  
 CC EMBL; X13060; CAA31458.1; --  
 CC PIR; A05139; A05139.  
 CC SWISS-2DPAGE; P02768; 1E7B.  
 CC MGD; MG1:87991; Alb1.  
 CC InterPro; IPR000264; Serum\_albumin.  
 CC Pfam; PF00273; transport\_prot; 3.  
 CC PRINTS; PR00802; SERUMALBUMIN.  
 CC ProDom; PD002486; Serum\_albumin; 1.  
 CC SMART; SM00103; ALBUMIN; 3.  
 CC PROSITE; PS00212; ALBUMIN; 3.  
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 CC SIGNAL 1 18  
 CC PROPEP 19 24  
 CC CHAIN 25 608  
 CC DOMAIN 25 205  
 CC DOMAIN 212 397  
 CC DOMAIN 404 595  
 CC METAL 27 86  
 CC DISULFID 77 86  
 CC DISULFID 99 115  
 CC DISULFID 114 125  
 CC DISULFID 148 193  
 CC DISULFID 192 201  
 CC DISULFID 224 270  
 CC DISULFID 269 277  
 CC DISULFID 289 303  
 CC DISULFID 302 313  
 CC DISULFID 340 385  
 CC DISULFID 384 393  
 CC DISULFID 416 462  
 CC DISULFID 461 472  
 CC DISULFID 485 501  
 CC DISULFID 538 583  
 CC CONFLICT 33 33  
 CC CONFLICT 41 41  
 CC CONFLICT 41 41  
 CC H -> D (IN REF. 6).  
 CC Q -> I (IN REF. 6).  
 CC Q -> I (IN REF. 6).  
 CC SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

Db 271 HGDLLLE 276

# RESULT 12

ALBU RABIT STANDARD; PRT; 608 AA.

AC P49065;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor.  
 GN ALB.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Liver;  
 RA Sheffield W.P., Syed S., Schuyler P.D.;  
 RL Submitted (DSC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U18344; AB558347.1; --  
 CC HSP; P02768; 1E7B.  
 CC InterPro; IPR000264; Serum\_albumin.  
 CC Pfam; PF00273; transport\_prot; 3.  
 CC PRINTS; PR00802; SERUMALBUMIN.  
 CC ProDom; PD002486; Serum\_albumin; 1.  
 CC SMART; SM00103; ALBUMIN; 3.  
 CC PROSITE; PS00212; ALBUMIN; 3.  
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 CC SIGNAL 1 18  
 CC PROPEP 19 24  
 CC CHAIN 25 608  
 CC DOMAIN 25 205  
 CC DOMAIN 212 397  
 CC DOMAIN 404 595  
 CC METAL 27 86  
 CC DISULFID 77 86  
 CC DISULFID 99 115  
 CC DISULFID 114 125  
 CC DISULFID 148 193  
 CC DISULFID 192 201  
 CC DISULFID 224 270  
 CC DISULFID 269 277  
 CC DISULFID 289 303  
 CC DISULFID 302 313  
 CC DISULFID 340 385  
 CC DISULFID 384 393  
 CC DISULFID 416 462  
 CC DISULFID 461 472  
 CC DISULFID 485 501  
 CC DISULFID 538 583  
 CC DISULFID 582 591  
 CC SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

Db 271 HGDLLLE 276

RESULT 13

ALBU\_RAT  
ID ALBU\_RAT STANDARD; PRT; 608 AA.  
AC P02770; P11382;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].  
GN ALB.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=91223722; PubMed=7017712;  
RX Sargent T.D., Yang M., Bonner J.;  
RA "Nucleotide sequence of cloned rat serum albumin messenger RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).  
[2]  
RN SEQUENCE OF 1-38, AND PROCESSING.  
RP MEDLINE=77249657; PubMed=893447;  
RX Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
RA "Rat liver pre-proalbumin: complete amino acid sequence of the pre-  
piece. Analysis of the direct translation product of albumin  
messenger RNA.";  
RL J. Biol. Chem. 252:6846-6855(1977).  
[3]  
RN SEQUENCE OF 25-222.  
RP MEDLINE=78109429; PubMed=564345;  
RX Isemura S., Ikenaka T.;  
RA "Amino acid sequences of fragments I and II obtained by cyanogen  
bromide cleavage of rat serum albumin.";  
RL J. Biochem. 83:35-48(1978).  
[4]  
RN SEQUENCE OF 223-288 AND 572-608.  
RP MEDLINE=76260153; PubMed=956149;  
RX Isemura S., Ikenaka T.;  
RA "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
the amino acid sequences of four fragments.";  
RL J. Biochem. 79:1183-1196(1976).  
[5]  
RN SEQUENCE OF 166-174.  
RP TISSUE=Plasma;  
RX MEDLINE=87194805; PubMed=2437111;  
RA Caraway R.E., Mitra S.P., Cochrane D.E.;  
RT "Structure of a biologically active neurotensin-related peptide  
obtained from pepsin-treated albumin(S).";  
RL J. Biol. Chem. 262:5968-5973(1987).  
[6]  
RN COPPER-BINDING.  
RX MEDLINE=73001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
RL Cancer Res. 38:3483-3486(1978).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and  
blood flow (potential).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; V01222; CAA24532.1; --  
DR PIR; A93872; ABETS.  
DR HSRF; P02768; 157B.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608  
FT PEPTIDE 166 174  
FT DOMAIN 25 205  
FT DOMAIN 212 397  
FT DOMAIN 404 595  
FT METAL 27 27  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 340 385  
FT DISULFID 384 393  
FT DISULFID 416 462  
FT DISULFID 461 472  
FT DISULFID 485 501  
FT DISULFID 500 511  
FT DISULFID 538 583  
FT DISULFID 582 591  
FT VARIANT 262 262  
FT CONFLICT 174 174  
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;  
Query Match 100.0%; Score 33; DB 1; Length 608;  
Best Local Similarity 100.0%; Pred. No. 16; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;  
OY 1 HGDLL 6  
DB 271 HGDLL 276  
RESULT 14  
ID ALBU\_HUMAN STANDARD; PRT; 609 AA.  
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=86196112; PubMed=3009475;  
RX Minghetti P.P., Ruffer D.E., Kuang W.J., Dennison O.E., Hawkins J.W.;  
RA Beattie W.G., Dugaiczyk A.;  
RT "Molecular structure of the human albumin gene is revealed by  
nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
[2]  
RN SEQUENCE FROM N.A.; AND VARIANT LYS-420.  
RP MEDLINE=82081882; PubMed=6171778;  
RX Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,

```

RESULT 8
Q941Y7
ID Q941Y7 PRELIMINARY; PRT; 280 AA.
AC Q941Y7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE P0431G06.23 protein (B1139B11.1 protein).
GN P0431G06.23 OR B1139B11.1
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530, 39947;
[1]
RN NCBI_TaxID=4530, 39947;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0431G06.23"
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone: B1139B11.1"
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AP003683; BAB64711.1; -
DR EMBL; AP004368; BAB90773.1; -
DR Gramene; Q941Y7; -
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 280 AA; 30370 MW; 097DD3165D61C76 CRC64;

Query Match 90.68; Score 29; DB 10; Length 280;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Q941Y7
Db 241 EADKAA 247

RESULT 9
Q8U6W5
ID Q8U6W5 PRELIMINARY; PRT; 311 AA.
AC Q8U6W5;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE 2-hydroxyacid dehydrogenase.
GN ATU4691 OR AGR_L_379.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavina T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,

```

```

RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323 (2001).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmle K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328 (2001).
DR EMBL; AE009397; AAL45485.1; -
DR EMBL; AE008217; AAK8759.1; -
DR PIR; AG3133; AG3133.
DR PIR; E98154; E98154.
DR GO; GO:001616; F:oxidoreductase activity, acting on the CH-O...; IEA.
DR GO; GO:0006564; P:L-serine biosynthesis; IEA.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH; 1.
DR Pfam; PF02826; 2-Hacid_DH_C; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 311 AA; 33047 MW; B6C98D520CC962EC CRC64;

Query Match 90.6%; Score 29; DB 16; Length 311;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Q941Y7
Db 30 EADKAA 36

RESULT 10
Q930D0
ID Q930D0 PRELIMINARY; PRT; 315 AA.
AC Q930D0;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative D-isomer specific 2-hydroxyacid.
GN RA0268 OR SMA0510.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=1021.
RC MEDLINE=21396509; PubMed=11481432;
RA Barlett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barclay-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
DR EMBL; AS007220; AAK64926.1; -
DR PIR; D95295; D95295.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:001616; F:oxidoreductase activity, acting on the CH-O...; IEA.
DR GO; GO:0006564; P:L-serine biosynthesis; IEA.
DR InterPro; IPR006139; 2-Hacid_DH.

```

OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33913 / NCPPB 528;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach P.C., Parah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorino C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Cimarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A.A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spinola L.A.P., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Secubal J.C., Kitajima J.P.;  
 RA "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AB012229; AAM40593.1; --  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 DR Hypothetical protein; Complete proteome.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 111 AA; 11336 MW; D2301B0058BC5445 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 111;  
 Best Local Similarity 85.7%; Pred. No. 1e+02; Indels 0; Gaps 0;  
 Matches 6; Conservative 1; Mismatches 0;

Cy 1 QAADKAA 7  
 Db 47 QAADRAA 53

RESULT 6

O87592 ID O87592 PRELIMINARY; PRT; 196 AA.  
 AC O87592;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein SC01650.  
 GN SC01650 OR SC14133C.  
 CN Streptomyces coelicolor.  
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2).  
 RA Nagy I., Tamura T., Vanderleyden J., Baumeister W., de Mot R.;  
 RT "The 20S proteasome of Streptomyces coelicolor";  
 RL J. Bacteriol. 190:0-0(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2).  
 RA Seeger K.J., Harris D.;  
 RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);

RX MEDLINE=97000351; PubMed=8943436;  
 RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Batsman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2)";  
 RL Nature 417:141-147(2002).  
 DR EMBL; AF086832; AAC64274.1; --  
 DR EMBL; AL939109; CAB59503.1; --  
 DR Hypothetical protein; Complete proteome.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 196 AA; 20239 MW; 35E647F0EB6256C4 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 196;  
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 QAADKAA 7  
 Db 45 QSAADKAA 51

RESULT 7

O7XPL3 ID O7XPL3 PRELIMINARY; PRT; 274 AA.  
 AC O7XPL3;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE OSJNB0085110.13 protein.  
 GN OSJNB0085110.13.  
 CN Oryza sativa (Rice).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,  
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL606684; CAE03568.1; --  
 SQ SEQUENCE 274 AA; 28866 MW; 499A949503543E5B CRC64;

Query Match 90.6%; Score 29; DB 10; Length 274;  
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 QAADKAA 7  
 Db 137 EAADKAA 143

```

Db      450 QAADKAA 456

RESULT 2
Q8YJB3  PRELIMINARY;      PRT;      96 AA.
ID      Q8YJB3;
AC      Q8YJB3;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Hypothetical cytosolic protein BMEI0173.
GN      BMEI0173.
OS      Brucella melitensis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29459;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=16M / ATCC 23456 / Biotype 1;
RX      MEDLINE=20020109; PubMed=11756888;
RA      DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujer C., Los T.,
RA      Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA      Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA      Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA      Haselkorn R., Kyripides N., Overbeek R.;
RT      "The genome sequence of the facultative intracellular pathogen
RT      Brucella melitensis.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR      EMBL; AE009459; AAL51355.1; -;
DR      PIR; AH3273; AH3273.
DR      InterPro; IPR005545; YCII.
DR      Pfam; PF03795; YCII; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 96 AA; 10359 MW; 0A876F82BCB68F6 CRC64;

Query Match      90.6%; Score 29; DB 16; Length 96;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      56 EAADKAA 62

RESULT 3
Q8FYI4  PRELIMINARY;      PRT;      96 AA.
ID      Q8FYI4;
AC      Q8FYI4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Conserved hypothetical protein.
GN      BR1890.
OS      Brucella suis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29461;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=1330 / Bivvar 1;
RX      MEDLINE=22247741; PubMed=12271122;
RA      Paulsen I.T., Seehadi R., Nelson K.E., Eisen J.A., Heidelberg J.P.,
RA      Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
RA      Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA      Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA      Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA      Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT      "The Brucella suis genome reveals fundamental similarities between
RT      animal and plant pathogens and symbionts.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR      EMBL; AE014479; AAN30784.1; -;
DR      PIR; AH3273; AH3273.
DR      TIGR; BR1890; -;

Query Match      90.6%; Score 29; DB 16; Length 108;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      45 QAADRAA 51

RESULT 5
Q8PB31  PRELIMINARY;      PRT;      111 AA.
ID      Q8PB31;
AC      Q8PB31;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein XCCI295.
GN      XCCI295.
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OX      NCBI_TaxID=92829;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=306 / ATCC 13902 / XV 101;
RX      MEDLINE=22022145; PubMed=12024217;
RA      da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA      Canarotte G., Cannavan F., Cardozo J., Chamberg J., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA      Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA      Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA      Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA      Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA      Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA      Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA      Setubal J.C., Kitajima J.P.;
RT      "Comparison of the genomes of two Xanthomonas pathogens with differing
RT      host specificities.";
RL      Nature 417:459-463(2002).
DR      EMBL; AE011766; AAM36217.1; -;
DR      InterPro; IPR000437; Prok_lipo_prot_S.
DR      PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 108 AA; 10946 MW; 9C20BAFDC8574728 CRC64;

Query Match      90.6%; Score 29; DB 16; Length 108;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      45 QAADRAA 51

RESULT 5
Q8PB31  PRELIMINARY;      PRT;      111 AA.
ID      Q8PB31;
AC      Q8PB31;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein XCCI295.
GN      XCCI295.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.70175 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_170\_176  
Perfect score: 32  
Sequence: 1 QAADKAA 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriaph:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID  | Description        |
|------------|-------|-------------|--------|--------|--------------------|
| 1          | 32    | 100.0       | 667    | Q89BJ7 | Q89BJ7 bradyrhizob |
| 2          | 29    | 90.6        | 96     | Q8FYD8 | Q8FYD8 brucella me |
| 3          | 29    | 90.6        | 96     | Q8FYD8 | Q8FYD8 brucella su |
| 4          | 29    | 90.6        | 108    | Q8PMT0 | Q8PMT0 xanthomonas |
| 5          | 29    | 90.6        | 111    | Q8PBT3 | Q8PBT3 xanthomonas |
| 6          | 29    | 90.6        | 196    | Q87592 | Q87592 streptomyc  |
| 7          | 29    | 90.6        | 274    | Q7XPL3 | Q7XPL3 oryza sativ |
| 8          | 29    | 90.6        | 280    | Q941Y7 | Q941Y7 oryza sativ |
| 9          | 29    | 90.6        | 311    | Q8U6W5 | Q8U6W5 agrobacteri |
| 10         | 29    | 90.6        | 315    | Q930D0 | Q930D0 rhizobium m |
| 11         | 29    | 90.6        | 387    | Q7U0Z1 | Q7U0Z1 mycobacteri |
| 12         | 29    | 90.6        | 466    | Q82PD6 | Q82PD6 streptomyc  |
| 13         | 29    | 90.6        | 503    | Q8NST2 | Q8NST2 corynebacte |
| 14         | 29    | 90.6        | 574    | Q8SXW8 | Q8SXW8 drosophila  |
| 15         | 29    | 90.6        | 584    | Q7VSG3 | Q7VSG3 felis silve |
| 16         | 29    | 90.6        | 607    | Q9W3Z7 | Q9W3Z7 drosophila  |

|    |    |      |      |    |        |                    |
|----|----|------|------|----|--------|--------------------|
| 17 | 29 | 90.6 | 906  | 16 | Q8YJG2 | Q8YJG2 brucella me |
| 18 | 29 | 90.6 | 906  | 16 | Q8FYD8 | Q8FYD8 brucella su |
| 19 | 29 | 90.6 | 1313 | 16 | Q8XRR8 | Q8XRR8 talstonia s |
| 20 | 29 | 90.6 | 1713 | 16 | Q82V52 | Q82V52 nitrosomona |
| 21 | 29 | 90.6 | 2215 | 5  | Q8II26 | Q8II26 plasmodium  |
| 22 | 28 | 87.5 | 120  | 16 | Q8F6P8 | Q8F6P8 leptospira  |
| 23 | 28 | 87.5 | 166  | 2  | Q9KJG2 | Q9KJG2 porphyromon |
| 24 | 28 | 87.5 | 210  | 10 | Q7XZX2 | Q7XZX2 oryza sativ |
| 25 | 28 | 87.5 | 225  | 17 | Q8TUH8 | Q8TUH8 methanosaer |
| 26 | 28 | 87.5 | 251  | 4  | Q8NSC7 | Q8NSC7 homo sapien |
| 27 | 28 | 87.5 | 260  | 16 | Q911L3 | Q911L3 pseudomonas |
| 28 | 28 | 87.5 | 260  | 16 | Q87L38 | Q87L38 vibrio para |
| 29 | 28 | 87.5 | 293  | 16 | Q9A2I9 | Q9A2I9 caulobacter |
| 30 | 28 | 87.5 | 310  | 16 | Q8A2T3 | Q8A2T3 bacteroides |
| 31 | 28 | 87.5 | 320  | 16 | Q9K3G9 | Q9K3G9 streptomyc  |
| 32 | 28 | 87.5 | 325  | 6  | Q28621 | Q28621 oryctolagus |
| 33 | 28 | 87.5 | 325  | 5  | Q26715 | Q26715 trypanosoma |
| 34 | 28 | 87.5 | 360  | 6  | Q9K6G9 | Q9K6G9 oryctolagus |
| 35 | 28 | 87.5 | 371  | 16 | Q7UYV0 | Q7UYV0 rhodopirell |
| 36 | 28 | 87.5 | 402  | 6  | Q28622 | Q28622 oryctolagus |
| 37 | 28 | 87.5 | 457  | 16 | Q7WQX0 | Q7WQX0 bordetella  |
| 38 | 28 | 87.5 | 457  | 16 | Q7W1Z2 | Q7W1Z2 bordetella  |
| 39 | 28 | 87.5 | 457  | 16 | Q7VUM3 | Q7VUM3 bordetella  |
| 40 | 28 | 87.5 | 462  | 16 | Q9CHJ4 | Q9CHJ4 lactococcus |
| 41 | 28 | 87.5 | 467  | 16 | Q8EDJ0 | Q8EDJ0 shevanelia  |
| 42 | 28 | 87.5 | 478  | 2  | Q846S1 | Q846S1 myxococcus  |
| 43 | 28 | 87.5 | 508  | 5  | Q9N4X2 | Q9N4X2 caenorhabdi |
| 44 | 28 | 87.5 | 520  | 5  | Q9BJF4 | Q9BJF4 toxoplasma  |
| 45 | 28 | 87.5 | 522  | 5  | Q26770 | Q26770 trypanosoma |

ALIGNMENTS

RESULT 1

Q89BJ7 PRELIMINARY; PRT; 667 AA.  
AC Q89BJ7;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE BLR8151 protein.  
GN BLR8151.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
RA Sasamoto S., Watanabe A., Ideasa K., Iizuchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RT Bradyrhizobium japonicum USDA110.";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AP005964; BAC53416.1; "  
DR CO; GO:0030693; F:casease activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001309; ICE\_P20.  
DR Pfam; PF00656; Peptidase C14; 1.  
DR PROSITE; PS0208; CASPASE\_P20; 1.  
KW Complete proteome.  
SQ SEQUENCE 667 AA; 70854 MW; 1394315BD73101C3 CRC64;

Query Match 100.0%; Score 32; DB 16; Length 667;

Best Local Similarity 100.0%; Pred.No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7

|||||



Mon Apr 19 13:27:12 2004

```

RESULT 15
TTF1_CAVPO
ID TTF1_CAVPO STANDARD; PRT; 112 AA.
AC P97273;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thyroid transcription factor 1 (Thyroid nuclear factor 1) (TTF-1)
DE (Homeobox protein Nkx-2.1) (Fragment).
GN TTF1 OR TTF1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunkin-Hartley; TISSUE=Lung;
RA Yuan H.T., Bingle C.D.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Transcription factor that binds and activates the
CC promoter of thyroid specific genes such as thyroglobulin,
CC thyroperoxidase, and thyrotropin receptor. Crucial in the
CC maintenance of the thyroid differentiation phenotype. May play a
CC role in lung development and surfactant homeostasis (By
CC similarity).
CC LOCATION: Nuclear.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the NK-2 homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U82718; AAB40921.1; -.
CC HSSP; P23441; 1FTT.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Transcription regulation; Activator; Homeobox; DNA-binding;
CC Nuclear protein.
CC KW NON_TER 1
CC FT DNA_BIND 1 60 HOMEBOX.
CC FT NON_TER 112 112
CC SQ SEQUENCE 112 AA; 12723 MW; AEAEDF06905F9DB CRC64;

Query Match 84.4%; Score 27; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QADKAA 7
Db 59 QAKDKAA 65

```

Search completed: April 19, 2004, 11:52:50  
Job time : 1.51062 secs

KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Multifunctional enzyme; FMN; FAD; Flavoprotein; NADP;  
 KW Complete proteome. 474  
 FT DOMAIN 1 CYTOCHROME P450.  
 FT DOMAIN 475 NADPH-P-450 REDUCTASE.  
 FT DOMAIN 493 FLAVOXYDOLIN-LIKE.  
 FT DOMAIN 666 880 FAD-BINDING.  
 FT METAL 403 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 1061 AA; 119467 MW; 7915DACC20578978 CRC64;

Query Match 87.5%; Score 28; DB 1; Length 1061;  
 Best Local Similarity 85.7%; Pred No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QAAKAA 7  
 DB 464 QAAKAA 470

RESULT 13  
 DEXT\_STRDO STANDARD; PRT; 1337 AA.  
 AC P39653;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-  
 DE glucanohydrolase).  
 GN DEX.  
 OS Streptococcus downei (Streptococcus sobrinus).  
 OG Plasmid pYA902.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus  
 OC NCBI\_TaxID=1317;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.  
 RC STRAIN=6715 / UAB66;  
 RX MEDLINE=94292401; PubMed=8021165;  
 RA Wanda S.-Y., Curtiss R. III;  
 RT "Purification and characterization of Streptococcus sobrinus  
 RT dextranase produced in recombinant Escherichia coli and sequence  
 RT analysis of the dextranase gene.";  
 RL J. Bacteriol. 176:3839-3850(1994).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE  
 CC PELICLE-COATED TOOTH SURFACE.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic  
 CC linkages in dextran.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (Potential).  
 CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3  
 CC and at 39 degrees Celsius.  
 CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.  
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M96978; AAA21772.1; -  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR TIGRfams; TIGR01167; LPXTG\_anchor; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;  
 KW Signal; Plasmid.  
 FT SIGNAL 1 30  
 FT CHAIN 31 1308 DEXTRANASE.  
 FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).  
 -----

Query Match 84.4%; Score 27; DB 1; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AADKAA 7  
 DB 22 AADKAA 27

FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).  
 FT MOD\_RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).  
 SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;  
 Query Match 87.5%; Score 28; DB 1; Length 1337;  
 Best Local Similarity 100.0%; Pred No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QAAKAA 6  
 DB 968 QAAKAA 973

RESULT 14  
 CHIB\_ARAHY STANDARD; PRT; 46 AA.  
 AC Q06013;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Endochitinase IB (EC 3.2.1.14) (CHIT 1B) (Fragment).  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OC NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91094785; PubMed=1980004;  
 RA Herget T., Schell J., Schreier P.H.;  
 RT "Elicitor-specific induction of one member of the chitinase gene  
 RT family in Arachis hypogaea.";  
 RL Mol. Gen. Genet. 224:469-476(1990).  
 CC -1- FUNCTION: This protein functions as a defense against chitin  
 CC containing fungal and bacterial pathogens.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
 CC acetyl-D-glucosamine polymers of chitin.  
 CC -1- INDUCTION: By yeast extract and dilution. Slight induction by  
 CC glucan elicitor.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS IA OR IB (BELONGS TO FAMILY  
 CC 19 OF GLYCOSYL HYDROLASES).  
 -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; X56891; CAA40210.1; -  
 DR HSSP; P23951; 2BAA.  
 DR InterPro; IPR000726; Glyco\_hydro\_19.  
 DR Pfam; PF00182; Glyco\_hydro\_19; 1.  
 DR PRODOM; PD354900; Glyco\_hydro\_19; 1.  
 DR PROSITE; PS00773; CHITINASE\_19\_1; PARTIAL.  
 DR PROSITE; PS00774; CHITINASE\_19\_2; PARTIAL.  
 KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding;  
 KW Multigene family.  
 FT NON\_TER 1 46  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 4672 MW; 4F868717729C2B9E CRC64;

Query Match 84.4%; Score 27; DB 1; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

protein synthesis. Protects formylmethionyl-tRNA from spontaneous hydrolysis and promotes its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (Cyttoplasmic).

-1- SUBCELLULAR LOCATION: Cytoplasmic.

-1- SIMILARITY: Belongs to the IP-2 family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; AF017154; AAP96265.1; -;  
 DR HMAP; MF 00100; -; 1.  
 DR InterPro; IPR004161; EFTU D2.  
 DR InterPro; IPR000795; EF\_GTPbind.  
 DR InterPro; IPR000178; IP2.  
 DR InterPro; IPR008847; IP2\_N\_GTP.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU D2; 2.  
 DR Pfam; PF04760; IP2\_N; 1.  
 DR ProDom; PD186100; IP2; 1.  
 DR TIGRFAMs; TIGR00487; IP-2; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS01176; IP2; 1.  
 DR Initiation factor; Protein biosynthesis; GTP-binding;  
 KW Complete proteome.  
 FT DOMAIN 341 490 G-DOMAIN.  
 FT NP\_BIND 347 354 GTP (BY SIMILARITY).  
 FT NP\_BIND 394 398 GTP (BY SIMILARITY).  
 FT NP\_BIND 448 451 GTP (BY SIMILARITY).  
 SQ SEQUENCE 839 AA; 91846 MW; BDFP37B8B10461D CRC64;

Query Match 87.5%; Score 28; DB 1; Length 839;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQKAA 7  
 |||:|  
 65 QAAEKAA 71

RESULT 12

ID\_CYPD\_BACSU STANDARD; PRT; 1061 AA.  
 AC O08394;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Probable bifunctional P-450:NADPH-P450 reductase 1 [Includes:  
 DE Cytochrome P450 102 (EC 1.14.14.1); NADPH-cytochrome P450 reductase  
 DE (EC 1.6.2.4)]  
 GN CYPD OR CYP102A2 OR BSU07250.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97453479; PubMed=9308178;  
 RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,  
 RA Duesterhoeft A., Ehrlich S.D.;  
 RT "Sequence of the Bacillus subtilis genome region in the vicinity of  
 RT the lev operon reveals two new extracytoplasmic function RNA  
 RT polymerase sigma factors SigV and SigZ";  
 RL Microbiology 143:2933-2943 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;

-----  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azavedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Enrian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Gusepbi C., Guy B.J., Haga K., Haisch J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra F., Tognotti A.,  
 RA Viati A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis";  
 RL Nature 390:249-256 (1997).  
 CC -1- FUNCTION: FUNCTIONS AS A FATTY ACID MONOOXYGENASE. THE REDUCTASE  
 CC DOMAIN IS REQUIRED FOR ELECTRON TRANSFER FROM NADP TO CYTOCHROME  
 CC P450 (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: NADPH + 2 ferricytochrome = NADP(+) + 2  
 CC ferrocycytochrome.  
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -1- COFACTOR: FAD and FMN (By similarity).  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the cytochrome  
 CC P450 family.  
 CC -1- SIMILARITY: Contains 1 FAD-binding domain.  
 CC -1- SIMILARITY: Contains 1 flavodoxin-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D87979; BAA20123.1; -;  
 DR EMBL; Z99107; CAB12544.1; -;  
 DR PIR; D69799; D69799.  
 DR HSSP; P14779; 1JPZ.  
 DR Subtilist; BG12871; cypD.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR003097; FAD\_Binding.  
 DR InterPro; IPR008254; Flav\_nitox\_synth.  
 DR InterPro; IPR001094; Flavodoxin\_like.  
 DR InterPro; IPR001709; FPN\_cyt\_reductse.  
 DR InterPro; IPR001433; Oxred\_FAD/NAD(P).  
 DR Pfam; PF00667; FAD\_binding\_1; 1.  
 DR Pfam; PF00358; flavodoxin; 1.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00369; FLAVODOXIN.  
 DR PRINTS; PR00371; FPNCR.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PROSITE; PS00902; FLAVODOXIN\_LIKE; 1.

STCA OR PKST.

OS *Emicella nidulans* (*Aspergillus nidulans*).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

CC Eurotiiales; Trichocomaceae; *Emicella*.

OX NCBI\_TaxID=162425;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=FGSC 26;

RC MEDLINE=96202293; PubMed=8643646;

RA Brown D.W., Yu J.-H., Kelkar H.S., Fernandes M., Neebitt T.C.,

RA Keller N.P., Adams T.H., Leonard T.J.;

RT "Twenty-five coregulated transcripts define a sterigmatocystin gene

RT cluster in *Aspergillus nidulans*."

RL Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422 (1996).

[2]

RN SEQUENCE FROM N.A.

RP STRAIN=FGSC 4;

RC MEDLINE=95370159; PubMed=7642507;

RA Yu J.-H., Leonard T.J.;

RT "Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a

RT novel type I polyketide synthase."

RL J. Bacteriol. 177:4792-4800 (1995).

CC -!- FUNCTION: Involved in the synthesis of the polyketide nucleus of

CC sterigmatocystin from hexanoyl-CoA and seven malonates.

CC -!- COFACTOR: Contains 2 covalently bound phosphopantetheines

CC (potential).

CC -!- PATHWAY: Sterigmatocystin biosynthesis; first step.

CC -!- SIMILARITY: Contains 2 acyl carrier domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U34740; AAC4191.1; --

DR EMBL; L39121; AAA81586.1; --

DR HSSP; P25715; IMLA.

DR InterPro; IPR008918; 5\_3\_exo\_C.

DR InterPro; IPR001227; AC trans.

DR InterPro; IPR000794; ketoacyl synth.

DR InterPro; IPR006163; Pp bind synth.

DR InterPro; IPR006162; Pp trans.

DR InterPro; IPR001031; thioesterase.

DR Pfam; PF00698; Acyl\_transf; 1.

DR Pfam; PF00109; ketoacyl-synt; 1.

DR Pfam; PF02801; ketoacyl-synt\_C; 1.

DR Pfam; PF00550; pp-binding; 2.

DR Pfam; PF00975; Thioesterase; 1.

DR PROSITE; PS00075; ACP DOMAIN; 2.

DR PROSITE; PS00012; PHOSPHOPANTHEINE; FALSE NEG.

KW Transferase; Acyltransferase; Phosphopantetheine; Repeat;

KW Multifunctional enzyme.

FT DOMAIN 383 814 BETA-KETOACYL SYNTHASE.

FT DOMAIN 884 1209 ACYL/MALONYL TRANSFERASES.

FT DOMAIN 1706 1777 ACYL CARRIER (ACP) 1.

FT DOMAIN 1830 1901 ACYL CARRIER (ACP) 2.

FT DOMAIN ? 2181 THIOESTERASE.

FT ACT\_SITE 552 552 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

FT ACT\_SITE 978 978 ACYL/MALONYL TRANSFERASES (BY

FT SIMILARITY).

FT BINDING 1738 1738 PHOSPHOPANTHEINE (BY SIMILARITY).

FT BINDING 1862 1862 PHOSPHOPANTHEINE (BY SIMILARITY).

FT ACT\_SITE 2028 2028 THIOESTERASE (BY SIMILARITY).

SQ SEQUENCE 2181 AA; 238831 MW; 5A3E5712AA9AD942 CRC64;

Query Match 90.6%; Score 29; DB 1; Length 2181;

Best Local Similarity 85.7%; Pred. No. 1.9e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QAADKAA 7

Db 1637 QASDXAA 1643

|||||

RESULT 10

MOBA THIFE

ID MOBA\_THIFE STANDARD; PRT; 409 AA.

AC P22898;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 01-JUL-1993 (Rel. 26, Last annotation update)

DE MOBA protein (Fragment).

GN MOBA.

OS *Thiobacillus ferrooxidans*.

OG Plasmid pTF-PC2.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;

OC Acidithiobacillaceae; Acidithiobacillus.

OX NCBI\_TaxID=920;

[1]

RN SEQUENCE FROM N.A.

RP MEDLINE=93015664; PubMed=1400173;

RA Rohrer J., Rawlings D.E.;

RT "Sequence analysis and characterization of the mobilization region of

RT a broad-host-range plasmid, pTF-PC2, isolated from *Thiobacillus*

RT *ferrooxidans*."

RL J. Bacteriol. 174:6230-6237 (1992).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; M57717; AAA27389.1; --

DR EMBL; A43256; A43256.

DR InterPro; IPR005094; Relaxase.

DR Pfam; PF03432; Relaxase; 1.

KW Plasmid.

FT NON TER 409 409

SQ SEQUENCE 409 AA; 46836 MW; 27C6593202572FE2 CRC64;

Query Match 87.5%; Score 28; DB 1; Length 409;

Best Local Similarity 85.7%; Pred. No. 70;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QAADKAA 7

|||||

Db 306 QAADKAA 312

|||||

RESULT 11

IF2\_HAEDU

ID IF2\_HAEDU STANDARD; PRT; 839 AA.

AC Q7VLI2;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Translation initiation factor IF-2.

GN INF2 OR HD1461

OS *Haemophilus ducreyi*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; *Haemophilus*.

OX NCBI\_TaxID=730;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=35000HP / ATCC 700724;

RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,

RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;

RT "The complete genome sequence of *Haemophilus ducreyi*."

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: One of the essential components for the initiation of

```

"Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RT  Gene 169:295-296(1996).
RL  -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC  binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC  hormones, bilirubin and drugs. Its main function is the regulation
CC  of the colloidal osmotic pressure of blood.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- TISSUE SPECIFICITY: Plasma.
CC  -!- ALLERGEN: Causes an allergic reaction in human.
CC  -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC  -!- SIMILARITY: Contains 3 albumin domains.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X84842; CAA59279.1; -.
DR  PIR; JC4660; S57632.
DR  HSSP; P02768; 1E7B.
DR  InterPro; IPR000264; Serum_albumin.
DR  Pfam; PF00273; transprot_prot; 3.
DR  PRINTS; PR00802; SERUMALBUMIN.
DR  ProDom; PD002486; Serum_albumin; 1.
DR  SMART; SM00103; ALBUMIN; 3.
DR  PROSITE; PS00212; ALBUMIN; 3.
DR  Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
KW  SIGNAL 1 18
FT  PROPEP 19 24
FT  CHAIN 25 608
FT  DOMAIN 25 205
FT  DOMAIN 212 397
FT  DOMAIN 404 595
FT  METAL 27 27
FT  DISULFID 77 86
FT  DISULFID 99 115
FT  DISULFID 114 125
FT  DISULFID 148 193
FT  DISULFID 192 201
FT  DISULFID 224 270
FT  DISULFID 269 277
FT  DISULFID 289 303
FT  DISULFID 302 313
FT  DISULFID 340 385
FT  DISULFID 384 393
FT  DISULFID 416 462
FT  DISULFID 461 472
FT  DISULFID 485 501
FT  DISULFID 500 511
FT  DISULFID 538 583
FT  DISULFID 582 591
SQ  SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;
Query Match 90.6%; Score 29; DB 1; Length 608;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0
OY 1 QAAAKAA 7
DB 194 EAADKAA 200
-----
RESULT 9
STCA_EMENI STANDARD; PRT; 2181 AA.
AC Q12397;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative sterigmatocystin biosynthesis polyketide synthase (PKS).

```

```

complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
MEDLINE=22206494; PubMed=12218036;
Fleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Keloney J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Beshai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
"Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
J. Bacteriol. 184:5479-5490(2002).
-!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
phosphate.
-!- PATHWAY: Tricarboxylic acid cycle.
-!- SUBUNIT: Composed of an alpha chain and a beta chain (By
similarity).
-!- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
subunit family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; Z79700; CAB01999.1; -.
EMBL; AE006983; AAK45226.1; -.
PIR; E70716; E70716.
HSP; P07460; 1SCU.
TIGR; MT0978; -.
Tuberculin; Rv0951; -.
HAMAP; MF 00558; 1.
InterPro; IPR001335; ATP-grasp.
InterPro; IPR005909; CoA_lig_beta.
InterPro; IPR005811; CoA_ligase.
Pfam; PF02222; ATP-grasp; 1.
Pfam; PF00549; ligase-CoA; 1.
TIGRFAMs; TIGR01016; succCoABeta; 1.
PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
Ligase; Tricarboxylic acid cycle; Complete proteome.
SEQUENCE 387 AA; 40925 MW; C3155A7CCAF98FDC CRC64;
Query Match 90.6%; Score 29; DB 1; Length 387;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0
QY 1 QAADKAA 7
Db 376 EAADKAA 382
-----
RESULT 8
ALBU_FELCA STANDARD; PRT; 608 AA.
ID ALBU_FELCA AC P490Z4;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
OS ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxId=9685;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;

```

```
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two albumins with identical electrophoretic mobility are produced
  Query Match 100.0%; Score 32; DB 1; Length 609;
  Best Local Similarity 100.0%; Pred. No. 13;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 194 QAADKAA 200
RESULT 5
RL7_STRV
ID RL7_STRV STANDARD; PRT; 126 AA.
AC P4936;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12 (SAL).
GN RPLL.
OS Streptomyces virginiae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1961;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96257210; PubMed=8675024;
RA Katayama M., Sakai Y., Okamoto S., Ihara F., Nihira T., Yamada Y.;
RT "Gene organization in the ada-rplL region of Streptomyces virginiae.";
RL Gene 171:135-136(1996).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
  involved in protein synthesis and appears to be essential for
  accurate translation (By similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; M89911; AAA26811.1; -
CC PIR; JCI273; JCI273.
CC HSP; P02392; ICTF.
CC HAMAP; MF 00368; -; 1.
CC InterPro; IPR008932; Ribos L12/7 olig.
CC InterPro; IPR000206; Ribosomal L12.
CC Pfam; PF00542; Ribosomal_L12; 1.
CC ProDom; PD001326; Ribosomal_L12; 1.
CC TIGRFAMs; TIGR00855; L12; 1.
CC TrEMBL; F5C3EE4F45D606E8 CRC64;
SQ SEQUENCE 126 AA; 13094 MW; 3A45F0754A65C74B CRC64;
  Query Match 90.6%; Score 29; DB 1; Length 126;
  Best Local Similarity 85.7%; Pred. No. 14;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 107 EAADKAA 113
RESULT 6
RL7_STRAT
ID RL7_STRAT STANDARD; PRT; 128 AA.
AC P29342;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92380478; PubMed=1511874;
RA Parra F., Blanco G., Alonso J.M., Balbin M., Mendez C., Salas J.A.;
RT "Cloning and sequence of a gene encoding the L7/L12 ribosomal protein
  equivalent of Streptomyces antibioticus.";
RL Gene 118:127-129(1992).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
  involved in protein synthesis and appears to be essential for
  accurate translation (By similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; M89911; AAA26811.1; -
CC PIR; JCI273; JCI273.
CC HSP; P02392; ICTF.
CC HAMAP; MF 00368; -; 1.
CC InterPro; IPR008932; Ribos L12/7 olig.
CC InterPro; IPR000206; Ribosomal L12.
CC Pfam; PF00542; Ribosomal_L12; 1.
CC ProDom; PD001326; Ribosomal_L12; 1.
CC TIGRFAMs; TIGR00855; L12; 1.
CC TrEMBL; F5C3EE4F45D606E8 CRC64;
SQ SEQUENCE 128 AA; 13272 MW; F5C3EE4F45D606E8 CRC64;
  Query Match 90.6%; Score 29; DB 1; Length 128;
  Best Local Similarity 85.7%; Pred. No. 15;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 109 EAADKAA 115
RESULT 7
SUCC_MYCTU
ID SUCC_MYCTU STANDARD; PRT; 387 AA.
AC P71559;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA synthetase beta chain (EC 6.2.1.5) (SCS-beta).
GN SUCC OR RV0951 OR MT0978 OR MTC10D7.23C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
```

Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

[7]  
 RT SEQUENCE FROM N.A. Skeletal muscle;  
 RN TISSUE=Liver, and Skeletal muscle;  
 RC MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smillius D.E.,  
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 25-609.  
 RX MEDLINE=76187907; PubMed=1225573;  
 RA Meloun B., Moravsek L., Kostka V.;  
 RT "Complete amino acid sequence of human serum albumin.";  
 RL FEBS Lett. 58:134-137(1975).  
 RN [9]  
 RP SEQUENCE OF 25-609.  
 RA Brown J.R., Shockley P., Behrens P.Q.;  
 RL (in) Bing D.H. (eds.);  
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
 RL Pergamon Press, New York (1979).  
 RN [10]  
 RP SEQUENCE OF 1-455 FROM N.A.  
 RC TISSUE=Liver;  
 RA Menaya J., Parrilla R., Ayuso M.S.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [11]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=86140099; PubMed=2419329;  
 RA Urano Y., Watanabe K., Sakai M., Tanaoki T.;  
 RT "The human albumin gene. Characterization of the 5' and 3' flanking  
 RT regions and the polymorphic gene transcripts.";  
 RL J. Biol. Chem. 261:3244-3251(1986).  
 RN [12]  
 RP SEQUENCE OF 222-229.  
 RX MEDLINE=76257808; PubMed=955075;  
 RA Walker J.E.;  
 RT "Lysine residue 199 of human serum albumin is modified by  
 RT acetylsalicylic acid.";  
 RL FEBS Lett. 66:173-175(1976).  
 RN [13]  
 RP SEQUENCE OF 25-44 AND 480-499.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [14]  
 RP DISULFIDE BONDS.  
 RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
 RT "Disulfide bonds in human serum albumin.";  
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
 RN [15]  
 RP BILIRUBIN-BINDING SITE.  
 RX MEDLINE=78186630; PubMed=656055;  
 RA Jacobsen C.;

"Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin";  
 Biochem. J. 171:453-459(1978).  
 RN [16]  
 RP VARIANT CANTERBURY ASN-337.  
 RX MEDLINE=87157744; PubMed=3828356;  
 RA Brennan S.O., Herbert P.;  
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
 RT domain of serum albumin.";  
 RL Biochim. Biophys. Acta 912:191-197(1987).  
 RN [17]  
 RP VARIANTS NAG-2 AND NAG-3.  
 RX MEDLINE=86088523; PubMed=3479777;  
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
 RA Satoh C., Neel J.V.;  
 RT "Amino acid substitutions in inherited albumin variants from  
 RT Amerindian and Japanese populations.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
 RN [18]  
 RP VARIANTS NAG-1, HIR-1, HIR-2 AND TOCHIGI.  
 RX MEDLINE=89345611; PubMed=2762316;  
 RA Arai K., Madison J., Hsu K., Ishioka N., Satoh C., Fujita M.,  
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
 RT "Point substitutions in Japanese alloalbumins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
 RN [19]  
 RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
 RX MEDLINE=90115905; PubMed=2404284;  
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
 RT "Point substitutions in albumin genetic variants from Asia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
 RN [20]  
 RP DESCRIPTION OF VARIANT REDHILL.  
 RX MEDLINE=90115852; PubMed=2104980;  
 RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;  
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
 RT human serum albumin whose precursor has an aberrant signal peptidase  
 RT cleavage site.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
 RN [21]  
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
 RX MEDLINE=91062352; PubMed=2247440;  
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
 RA Watkins S., Putnam F.W.;  
 RT "Mutations in genetic variants of human serum albumin found in  
 RT Italy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
 RN [22]  
 RP VARIANT VENEZIA.  
 RX MEDLINE=91296740; PubMed=2068071;  
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
 RA Minchiotti L., Putnam F.W.;  
 RT "A donor splice mutation and a single-base deletion produce two  
 RT carboxyl-terminal variants of human serum albumin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
 RN [23]  
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
 RX MEDLINE=92052189; PubMed=1946412;  
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
 RA Matsuda Y., Amaki I., Putnam F.W.;  
 RT "Genetic variants of serum albumin in Americans and Japanese.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
 RN [24]  
 RP VARIANT CASEBROOK ASN-518.  
 RX MEDLINE=91316157; PubMed=1859851;  
 RA Peach R.J., Brennan S.O.;  
 RT "Structural characterization of a glycoprotein variant of human serum  
 RT albumin: albumin Casebrook (494 Asp-->Asn).";  
 RL Biochim. Biophys. Acta 1097:49-54(1991).  
 RN [25]  
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
 RX MEDLINE=92190239; PubMed=1347703;

RT "Isolation, amino acid sequence and copper(II)-binding properties of  
 RL peptide (1-24) of dog serum albumin.";  
 RN J. Biol. Chem. 249:5872-5877(1974).  
 RN [4]  
 RP SEQUENCE OF 25-38.  
 RC TISSUE=Heart;  
 RA MEDLINE=9816340; PubMed=9504812;  
 RX Duan M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-DPAGE and the two-dimensional gel electrophoresis database of  
 RT dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 RN [5]  
 RP SEQUENCE OF 215-478 FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=94201492; PubMed=7512102;  
 RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,  
 RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
 RT "Molecular characterization of dog albumin as a cross-reactive  
 RT allergen.";  
 RL J. Allergy Clin. Immunol. 93:614-627(1994).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human.  
 CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ133489; CAB64867.1; -;  
 DR EMBL; Y17737; CA76841.1; -;  
 DR EMBL; S72946; AAB30434.1; -;  
 DR HSRP; P02768; I57B.  
 DR HSC-2DPAGE; P49822; DOG.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transprot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608 SERUM ALBUMIN.  
 FT DOMAIN 25 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.  
 FT DOMAIN 404 595 ALBUMIN 3.  
 FT METAL 27 27 COPPER (BY SIMILARITY).  
 FT DISULFID 77 86 BY SIMILARITY.  
 FT DISULFID 99 115 BY SIMILARITY.  
 FT DISULFID 114 125 BY SIMILARITY.  
 FT DISULFID 148 193 BY SIMILARITY.  
 FT DISULFID 192 201 BY SIMILARITY.  
 FT DISULFID 224 270 BY SIMILARITY.  
 FT DISULFID 269 277 BY SIMILARITY.  
 FT DISULFID 289 303 BY SIMILARITY.  
 FT DISULFID 302 313 BY SIMILARITY.  
 FT DISULFID 340 385 BY SIMILARITY.  
 FT DISULFID 384 393 BY SIMILARITY.  
 FT DISULFID 416 462 BY SIMILARITY.  
 FT DISULFID 461 472 BY SIMILARITY.  
 FT DISULFID 485 501 BY SIMILARITY.  
 FT DISULFID 500 511 BY SIMILARITY.  
 FT DISULFID 538 583 BY SIMILARITY.

FT DISULFID 582 591 BY SIMILARITY.  
 FT CONFLICT 1 26 MEKVTFFISLFFLFSSAYSGRLVREA -> MDT (IN  
 FT REF. 2).  
 FT CONFLICT 146 146 A -> R (IN REF. 2).  
 FT CONFLICT 206 206 I -> T (IN REF. 2).  
 FT CONFLICT 349 349 V -> A (IN REF. 2).  
 FT CONFLICT 359 359 S -> A (IN REF. 1).  
 FT CONFLICT 448 448 V -> VV (IN REF. 5).  
 FT CONFLICT 474 474 D -> E (IN REF. 1).  
 SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;  
 Query Match 100.0%; Score 32; DB 1; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QAAADKAA 7  
 Db 194 QAAADKAA 200  
 RESULT 4  
 ALBU HUMAN STANDARD; PRT: 609 AA.  
 ID P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;  
 DT 21-JUN-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OK NCBI\_taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86196112; PubMed=3009475;  
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
 RA Beattie W.G., Dugaiczky A.;  
 RT "Molecular structure of the human albumin gene is revealed by  
 RT nucleotide sequence within gil-22 of chromosome 4.";  
 RL J. Biol. Chem. 261:6747-6757(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
 RX MEDLINE=82081882; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.;  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli.";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
 RX MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.;  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (PRO0503/PRO1708/PRO2044/PRO2615/PRO2675).  
 RC TISSUE=Fetal liver;  
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
 RT "Functional prediction of the coding sequences of 121 new genes  
 RT deduced by analysis of cDNA clones from human fetal liver.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A., AND VARIANT HIPOSHIMA-1 LYS-378.  
 RA Huang M.C., Wu H.T.;  
 RT "The cDNA sequences of human serum albumin.";



```
FT METAL 19 19 COPPER (BY SIMILARITY)..
FT BINDING 256 256 BILIRUBIN (POTENTIAL)..
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 32; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 186 QAADKAA 192

RESULT 2
ALBU_PIG STANDARD; PRT; 605 AA.
ID ALBU_PIG
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S.; Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X12422; CAA30970.1; --
CC EMBL; M36787; AAA30988.1; --
CC PIR; S01382; ABPGS.
CC HSSP; P02768; 1E7H.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
```

```
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON TER 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match 100.0%; Score 32; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 191 QAADKAA 197

RESULT 3
ALBU_CANFA STANDARD; PRT; 608 AA.
ID ALBU_CANFA
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjatan B.; Svoboda I.; Brandejsky-Pichler F.; Rumpold H.;
RA Valenta R.; Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W.; Sarkar B.;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.510619 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_170\_176

Perfect score: 32

Sequence: 1 QAADKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1          | 32    | 100.0         | 600    | 1     | ALBU_MACMU  |
| 2          | 32    | 100.0         | 605    | 1     | ALBU_FIG    |
| 3          | 32    | 100.0         | 608    | 1     | ALBU_CANFA  |
| 4          | 32    | 100.0         | 609    | 1     | ALBU_HUMAN  |
| 5          | 29    | 90.6          | 126    | 1     | RL7_STRVG   |
| 6          | 29    | 90.6          | 128    | 1     | RL7_STRAT   |
| 7          | 29    | 90.6          | 387    | 1     | SUCC_MICTU  |
| 8          | 29    | 90.6          | 608    | 1     | ALBU_FELCA  |
| 9          | 29    | 90.6          | 2181   | 1     | STCA_EMENI  |
| 10         | 28    | 87.5          | 409    | 1     | MOBA_THIFE  |
| 11         | 28    | 87.5          | 839    | 1     | IF2_HARDU   |
| 12         | 28    | 87.5          | 1051   | 1     | CYPD_BACSU  |
| 13         | 28    | 87.5          | 1337   | 1     | DEXT_STRDO  |
| 14         | 27    | 84.4          | 46     | 1     | CHIB_ARAHY  |
| 15         | 27    | 84.4          | 112    | 1     | TTFI_CAVPO  |
| 16         | 27    | 84.4          | 129    | 1     | RS8_ECOLI   |
| 17         | 27    | 84.4          | 161    | 1     | HLPA_ECOLI  |
| 18         | 27    | 84.4          | 161    | 1     | OMPH_SALTY  |
| 19         | 27    | 84.4          | 165    | 1     | OMPH_VERPE  |
| 20         | 27    | 84.4          | 286    | 1     | PTNC_ECOLI  |
| 21         | 27    | 84.4          | 266    | 1     | PTRC_KLEPN  |
| 22         | 27    | 84.4          | 238    | 1     | HISI_VIECH  |
| 23         | 27    | 84.4          | 305    | 1     | ROAO_HUMAN  |
| 24         | 27    | 84.4          | 353    | 1     | NK24_HUMAN  |
| 25         | 27    | 84.4          | 354    | 1     | NC24_MOUSE  |
| 26         | 27    | 84.4          | 366    | 1     | LC05_HUMAN  |
| 27         | 27    | 84.4          | 371    | 1     | TTFI_CANFA  |
| 28         | 27    | 84.4          | 371    | 1     | TTFI_HUMAN  |
| 29         | 27    | 84.4          | 372    | 1     | TTFI_MOUSE  |
| 30         | 27    | 84.4          | 372    | 1     | TTFI_RAT    |
| 31         | 27    | 84.4          | 378    | 1     | MOBL_THIFE  |
| 32         | 27    | 84.4          | 384    | 1     | TMPB_TREPH  |
| 33         | 27    | 84.4          | 394    | 1     | FLBI_RHIME  |

34 84.4 394 1 SUC1\_STRCO  
35 27 84.4 395 1 FLB2\_RHIME  
36 27 84.4 500 1 YDGR\_ECOLI  
37 27 84.4 506 1 PCGB\_BACSU  
38 27 84.4 558 1 SCFN\_MOUSE  
39 27 84.4 559 1 SCFN\_HUMAN  
40 27 84.4 665 1 LAMA\_MOUSE  
41 27 84.4 665 1 LAMA\_RAT  
42 27 84.4 707 1 RHO\_STRLI  
43 27 84.4 904 1 NUOG\_PSEFL  
44 27 84.4 1321 1 MDRI\_CAEEL  
45 27 84.4 3644 1 MINT\_MOUSE

#### ALIGNMENTS

RESULT 1  
ID \_ALBU\_MACMU STANDARD; PRT; 600 AA.  
AC Q28522;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor (fragment).  
GN ALB  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
CX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9321191; PubMed=8460152;  
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,  
RA Dwulet J., Putnam F.W.;  
RT "cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC EMBL; M90463; AAA36906.1; -;  
CC PIR; A47391; A47391.  
CC HSSP; P02768; 1E7B.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC PRODOM; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT NON TER 1  
FT SIGNAL <1 10 BY SIMILARITY.  
FT PROPEP 11 16 BY SIMILARITY.  
FT CHAIN 17 600 SERUM ALBUMIN.  
FT DOMAIN 17 197 ALBUMIN 1.  
FT DOMAIN 204 389 ALBUMIN 2.  
FT DOMAIN 396 587 ALBUMIN 3.

Mon Apr 19 13:27:27 2004

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 RN NCB1\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MR-1;  
 RX MEDLINE=22297686; PubMed=12368813;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
 RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,  
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,  
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealsen K.H., Fraser C.M.;  
 RA "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RT Shewanella oneidensis";  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL; AE015709; AAN5728.1; -.  
 DR TIGR; SO2700; -.  
 DR InterPro; IPR006531; Phage\_P2\_V.  
 DR TIGRFAMs; TIGR01644; phage\_P2\_V; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 215 AA; 23197 MW; 8ACFDA22355AD55E CRC64;  
 Query Match 81.6%; Score 31; DB 16; Length 215;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ADESAEN 8  
 Db |||||:  
 45 ADESDN 51

Search completed: April 19, 2004, 11:59:59  
 Job time : 5.08772 secs

```
Query Match      84.2%; Score 32; DB 5; Length 1128;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAN 8
Db 652 VADENACN 659
|||||:|:|

RESULT 12
Q8WQC8 PRELIMINARY; PRT; 1702 AA.
AC Q8WQC8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T13H2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wu X.;
RT "The sequence of C. elegans cosmid T13H2.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U39563; AAL56623.1; -
DR WormPep; T13H2.4a; C330179.
KW Hypothetical protein.
SQ SEQUENCE 1702 AA; 190617 MW; 1F5CAC942100FEB5 CRC64;

Query Match      84.2%; Score 32; DB 5; Length 1702;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAN 7
Db 129 VADESAN 135
|||||:|:|

RESULT 13
Q9F428 PRELIMINARY; PRT; 75 AA.
AC Q9F428;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE TEL1 protein.
GN TEL1.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO28;
RA Perez-Diaz J.C.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBSJ databases.
```

```
DR EMBL; AJ012114; CAC13962.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
SQ SEQUENCE 75 AA; 8192 MW; 5F1899B3C945B4BF CRC64;

Query Match      81.6%; Score 31; DB 2; Length 75;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAN 8
Db 66 VADDSVEN 73
|||||:|:|

RESULT 14
Q8Y784 PRELIMINARY; PRT; 75 AA.
AC Q8Y784;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein lmo1411.
GN lmo1411.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
Raquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
Charbit A., Chetoui F., Couve E., de Daruvar A., Denoux P.,
Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,
Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapat G.,
Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591979; CAC99489.1; -
DR PIR; AC1251; AC1251.
DR Listlist; LMO01411; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH_3.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8263 MW; 545899B3C549B4E3 CRC64;

Query Match      81.6%; Score 31; DB 16; Length 75;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAN 8
Db 66 VADDSVEN 73
|||||:|:|

RESULT 15
Q8EDP3 PRELIMINARY; PRT; 215 AA.
AC Q8EDP3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Prophage MuSo2, baseplate assembly protein V.
GN SO2700.
OS Shewanella oneidensis.
```

RT 60,770 full-length cDNAs. ;  
RL Nature 420:563-573 (2002). ;  
DR EMBL; AK050644; BAC34360.1; --  
DR MGD; MGI:87991; Albl.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR00264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
DR NON\_TER 1  
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;  
  
Query Match 84.2%; Score 32; DB 11; Length 576;  
Best Local Similarity 87.5%; Pred. No. 1.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 VADESAEN 8  
Db 46 VADESAEN 53  
|||||  
  
RESULT 9  
Q8C7H3 PRELIMINARY; PRT; 608 AA.  
ID Q8C7H3  
AC Q8C7H3  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Albumin 1.  
GN ALB1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs. ;  
RL Nature 420:563-573 (2002). ;  
DR EMBL; AK050644; BAC34145.1; --  
DR MGD; MGI:87991; Albl.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR00264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;  
  
Query Match 84.2%; Score 32; DB 11; Length 608;  
Best Local Similarity 87.5%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 VADESAEN 8  
Db 78 VADESAEN 85  
|||||  
  
RESULT 10  
Q8WQC7 PRELIMINARY; PRT; 649 AA.  
ID Q8WQC7  
AC Q8WQC7

DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN T13H2.4  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium. ;  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX Wu X.;  
RT "The sequence of C. elegans cosmid T13H2. ;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX Waterston R.;  
RT "Direct Submission. ;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U39653; AAL56624.1; --  
DR WormPep; T13H2.4b; CE30180.  
KW Hypothetical protein.  
SQ SEQUENCE 649 AA; 72873 MW; 598793DFF63A506 CRC64;  
  
Query Match 84.2%; Score 32; DB 5; Length 649;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VADESAE 7  
Db 129 VADESAE 135  
|||||  
  
RESULT 11  
O62235 PRELIMINARY; PRT; 1128 AA.  
ID O62235  
AC O62235  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein F36F2.3.  
GN F36F2.3  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
RT investigating biology. ;  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX Cottage A.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z81532; CAB04326.3; --  
KW Hypothetical protein.  
SQ SEQUENCE 1128 AA; 130372 MW; 6445EE26D7D4775D CRC64;

TIGR; TC0331; --  
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0016987; F:sigma factor activity; IEA.  
DR GO; GO:0003700; P:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR GO; GO:0006352; P:transcription initiation; IEA.  
DR InterPro; IPR009043; RNA\_pol\_sigma.  
DR InterPro; IPR007627; Sigma70\_r2.  
DR InterPro; IPR007624; Sigma70\_r3.  
DR InterPro; IPR007630; Sigma70\_r4.  
DR InterPro; IPR000943; Sigma70.  
DR Pfam; PF04542; sigma70\_r2; 1.  
DR Pfam; PF04539; sigma70\_r3; 1.  
DR Pfam; PF04545; sigma70\_r4; 1.  
DR PRINTS; PR00046; SIGMA70FC1.  
DR DNA-binding; DNA-directed RNA polymerase; Sigma factor;  
KW Transcription regulation; Complete proteome.  
SQ SEQUENCE 253 AA; 28962 MW; 8365F50D21929538 CRC64;  
  
Query Match 84.2%; Score 32; DB 16; Length 253;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 VADESAEN 8  
Db 174 IADERAEN 181  
:|||||  
[1]  
[2]  
[3]  
[4]  
[5]  
[6]  
[7]  
[8]  
[9]  
[10]  
[11]  
[12]  
[13]  
[14]  
[15]  
[16]  
[17]  
[18]  
[19]  
[20]  
[21]  
[22]  
[23]  
[24]  
[25]  
[26]  
[27]  
[28]  
[29]  
[30]  
[31]  
[32]  
[33]  
[34]  
[35]  
[36]  
[37]  
[38]  
[39]  
[40]  
[41]  
[42]  
[43]  
[44]  
[45]  
[46]  
[47]  
[48]  
[49]  
[50]  
[51]  
[52]  
[53]  
[54]  
[55]  
[56]  
[57]  
[58]  
[59]  
[60]  
[61]  
[62]  
[63]  
[64]  
[65]  
[66]  
[67]  
[68]  
[69]  
[70]  
[71]  
[72]  
[73]  
[74]  
[75]  
[76]  
[77]  
[78]  
[79]  
[80]  
[81]  
[82]  
[83]  
[84]  
[85]  
[86]  
[87]  
[88]  
[89]  
[90]  
[91]  
[92]  
[93]  
[94]  
[95]  
[96]  
[97]  
[98]  
[99]  
[100]  
[101]  
[102]  
[103]  
[104]  
[105]  
[106]  
[107]  
[108]  
[109]  
[110]  
[111]  
[112]  
[113]  
[114]  
[115]  
[116]  
[117]  
[118]  
[119]  
[120]  
[121]  
[122]  
[123]  
[124]  
[125]  
[126]  
[127]  
[128]  
[129]  
[130]  
[131]  
[132]  
[133]  
[134]  
[135]  
[136]  
[137]  
[138]  
[139]  
[140]  
[141]  
[142]  
[143]  
[144]  
[145]  
[146]  
[147]  
[148]  
[149]  
[150]  
[151]  
[152]  
[153]  
[154]  
[155]  
[156]  
[157]  
[158]  
[159]  
[160]  
[161]  
[162]  
[163]  
[164]  
[165]  
[166]  
[167]  
[168]  
[169]  
[170]  
[171]  
[172]  
[173]  
[174]  
[175]  
[176]  
[177]  
[178]  
[179]  
[180]  
[181]  
[182]  
[183]  
[184]  
[185]  
[186]  
[187]  
[188]  
[189]  
[190]  
[191]  
[192]  
[193]  
[194]  
[195]  
[196]  
[197]  
[198]  
[199]  
[200]  
[201]  
[202]  
[203]  
[204]  
[205]  
[206]  
[207]  
[208]  
[209]  
[210]  
[211]  
[212]  
[213]  
[214]  
[215]  
[216]  
[217]  
[218]  
[219]  
[220]  
[221]  
[222]  
[223]  
[224]  
[225]  
[226]  
[227]  
[228]  
[229]  
[230]  
[231]  
[232]  
[233]  
[234]  
[235]  
[236]  
[237]  
[238]  
[239]  
[240]  
[241]  
[242]  
[243]  
[244]  
[245]  
[246]  
[247]  
[248]  
[249]  
[250]  
[251]  
[252]  
[253]  
[254]  
[255]  
[256]  
[257]  
[258]  
[259]  
[260]  
[261]  
[262]  
[263]  
[264]  
[265]  
[266]  
[267]  
[268]  
[269]  
[270]  
[271]  
[272]  
[273]  
[274]  
[275]  
[276]  
[277]  
[278]  
[279]  
[280]  
[281]  
[282]  
[283]  
[284]  
[285]  
[286]  
[287]  
[288]  
[289]  
[290]  
[291]  
[292]  
[293]  
[294]  
[295]  
[296]  
[297]  
[298]  
[299]  
[300]  
[301]  
[302]  
[303]  
[304]  
[305]  
[306]  
[307]  
[308]  
[309]  
[310]  
[311]  
[312]  
[313]  
[314]  
[315]  
[316]  
[317]  
[318]  
[319]  
[320]  
[321]  
[322]  
[323]  
[324]  
[325]  
[326]  
[327]  
[328]  
[329]  
[330]  
[331]  
[332]  
[333]  
[334]  
[335]  
[336]  
[337]  
[338]  
[339]  
[340]  
[341]  
[342]  
[343]  
[344]  
[345]  
[346]  
[347]  
[348]  
[349]  
[350]  
[351]  
[352]  
[353]  
[354]  
[355]  
[356]  
[357]  
[358]  
[359]  
[360]  
[361]  
[362]  
[363]  
[364]  
[365]  
[366]  
[367]  
[368]  
[369]  
[370]  
[371]  
[372]  
[373]  
[374]  
[375]  
[376]  
[377]  
[378]  
[379]  
[380]  
[381]  
[382]  
[383]  
[384]  
[385]  
[386]  
[387]  
[388]  
[389]  
[390]  
[391]  
[392]  
[393]  
[394]  
[395]  
[396]  
[397]  
[398]  
[399]  
[400]  
[401]  
[402]  
[403]  
[404]  
[405]  
[406]  
[407]  
[408]  
[409]  
[410]  
[411]  
[412]  
[413]  
[414]  
[415]  
[416]  
[417]  
[418]  
[419]  
[420]  
[421]  
[422]  
[423]  
[424]  
[425]  
[426]  
[427]  
[428]  
[429]  
[430]  
[431]  
[432]  
[433]  
[434]  
[435]  
[436]  
[437]  
[438]  
[439]  
[440]  
[441]  
[442]  
[443]  
[444]  
[445]  
[446]  
[447]  
[448]  
[449]  
[450]  
[451]  
[452]  
[453]  
[454]  
[455]  
[456]  
[457]  
[458]  
[459]  
[460]  
[461]  
[462]  
[463]  
[464]  
[465]  
[466]  
[467]  
[468]  
[469]  
[470]  
[471]  
[472]  
[473]  
[474]  
[475]  
[476]  
[477]  
[478]  
[479]  
[480]  
[481]  
[482]  
[483]  
[484]  
[485]  
[486]  
[487]  
[488]  
[489]  
[490]  
[491]  
[492]  
[493]  
[494]  
[495]  
[496]  
[497]  
[498]  
[499]  
[500]  
[501]  
[502]  
[503]  
[504]  
[505]  
[506]  
[507]  
[508]  
[509]  
[510]  
[511]  
[512]  
[513]  
[514]  
[515]  
[516]  
[517]  
[518]  
[519]  
[520]  
[521]  
[522]  
[523]  
[524]  
[525]  
[526]  
[527]  
[528]  
[529]  
[530]  
[531]  
[532]  
[533]  
[534]  
[535]  
[536]  
[537]  
[538]  
[539]  
[540]  
[541]  
[542]  
[543]  
[544]  
[545]  
[546]  
[547]  
[548]  
[549]  
[550]  
[551]  
[552]  
[553]  
[554]  
[555]  
[556]  
[557]  
[558]  
[559]  
[560]  
[561]  
[562]  
[563]  
[564]  
[565]  
[566]  
[567]  
[568]  
[569]  
[570]  
[571]  
[572]  
[573]  
[574]  
[575]  
[576]  
[577]  
[578]  
[579]  
[580]  
[581]  
[582]  
[583]  
[584]  
[585]  
[586]  
[587]  
[588]  
[589]  
[590]  
[591]  
[592]  
[593]  
[594]  
[595]  
[596]  
[597]  
[598]  
[599]  
[600]  
[601]  
[602]  
[603]  
[604]  
[605]  
[606]  
[607]  
[608]  
[609]  
[610]  
[611]  
[612]  
[613]  
[614]  
[615]  
[616]  
[617]  
[618]  
[619]  
[620]  
[621]  
[622]  
[623]  
[624]  
[625]  
[626]  
[627]  
[628]  
[629]  
[630]  
[631]  
[632]  
[633]  
[634]  
[635]  
[636]  
[637]  
[638]  
[639]  
[640]  
[641]  
[642]  
[643]  
[644]  
[645]  
[646]  
[647]  
[648]  
[649]  
[650]  
[651]  
[652]  
[653]  
[654]  
[655]  
[656]  
[657]  
[658]  
[659]  
[660]  
[661]  
[662]  
[663]  
[664]  
[665]  
[666]  
[667]  
[668]  
[669]  
[670]  
[671]  
[672]  
[673]  
[674]  
[675]  
[676]  
[677]  
[678]  
[679]  
[680]  
[681]  
[682]  
[683]  
[684]  
[685]  
[686]  
[687]  
[688]  
[689]  
[690]  
[691]  
[692]  
[693]  
[694]  
[695]  
[696]  
[697]  
[698]  
[699]  
[700]  
[701]  
[702]  
[703]  
[704]  
[705]  
[706]  
[707]  
[708]  
[709]  
[710]  
[711]  
[712]  
[713]  
[714]  
[715]  
[716]  
[717]  
[718]  
[719]  
[720]  
[721]  
[722]  
[723]  
[724]  
[725]  
[726]  
[727]  
[728]  
[729]  
[730]  
[731]  
[732]  
[733]  
[734]  
[735]  
[736]  
[737]  
[738]  
[739]  
[740]  
[741]  
[742]  
[743]  
[744]  
[745]  
[746]  
[747]  
[748]  
[749]  
[750]  
[751]  
[752]  
[753]  
[754]  
[755]  
[756]  
[757]  
[758]  
[759]  
[760]  
[761]  
[762]  
[763]  
[764]  
[765]  
[766]  
[767]  
[768]  
[769]  
[770]  
[771]  
[772]  
[773]  
[774]  
[775]  
[776]  
[777]  
[778]  
[779]  
[780]  
[781]  
[782]  
[783]  
[784]  
[785]  
[786]  
[787]  
[788]  
[789]  
[790]  
[791]  
[792]  
[793]  
[794]  
[795]  
[796]  
[797]  
[798]  
[799]  
[800]  
[801]  
[802]  
[803]  
[804]  
[805]  
[806]  
[807]  
[808]  
[809]  
[810]  
[811]  
[812]  
[813]  
[814]  
[815]  
[816]  
[817]  
[818]  
[819]  
[820]  
[821]  
[822]  
[823]  
[824]  
[825]  
[826]  
[827]  
[828]  
[829]  
[830]  
[831]  
[832]  
[833]  
[834]  
[835]  
[836]  
[837]  
[838]  
[839]  
[840]  
[841]  
[842]  
[843]  
[844]  
[845]  
[846]  
[847]  
[848]  
[849]  
[850]  
[851]  
[852]  
[853]  
[854]  
[855]  
[856]  
[857]  
[858]  
[859]  
[860]  
[861]  
[862]  
[863]  
[864]  
[865]  
[866]  
[867]  
[868]  
[869]  
[870]  
[871]  
[872]  
[873]  
[874]  
[875]  
[876]  
[877]  
[878]  
[879]  
[880]  
[881]  
[882]  
[883]  
[884]  
[885]  
[886]  
[887]  
[888]  
[889]  
[890]  
[891]  
[892]  
[893]  
[894]  
[895]  
[896]  
[897]  
[898]  
[899]  
[900]  
[901]  
[902]  
[903]  
[904]  
[905]  
[906]  
[907]  
[908]  
[909]  
[910]  
[911]  
[912]  
[913]  
[914]  
[915]  
[916]  
[917]  
[918]  
[919]  
[920]  
[921]  
[922]  
[923]  
[924]  
[925]  
[926]  
[927]  
[928]  
[929]  
[930]  
[931]  
[932]  
[933]  
[934]  
[935]  
[936]  
[937]  
[938]  
[939]  
[940]  
[941]  
[942]  
[943]  
[944]  
[945]  
[946]  
[947]  
[948]  
[949]  
[950]  
[951]  
[952]  
[953]  
[954]  
[955]  
[956]  
[957]  
[958]  
[959]  
[960]  
[961]  
[962]  
[963]  
[964]  
[965]  
[966]  
[967]  
[968]  
[969]  
[970]  
[971]  
[972]  
[973]  
[974]  
[975]  
[976]  
[977]  
[978]  
[979]  
[980]  
[981]  
[982]  
[983]  
[984]  
[985]  
[986]  
[987]  
[988]  
[989]  
[990]  
[991]  
[992]  
[993]  
[994]  
[995]  
[996]  
[997]  
[998]  
[999]  
[1000]  
[1001]  
[1002]  
[1003]  
[1004]  
[1005]  
[1006]  
[1007]  
[1008]  
[1009]  
[1010]  
[1011]  
[1012]  
[1013]  
[1014]  
[1015]  
[1016]  
[1017]  
[1018]  
[1019]  
[1020]  
[1021]  
[1022]  
[1023]  
[1024]  
[1025]  
[1026]  
[1027]  
[1028]  
[1029]  
[1030]  
[1031]  
[1032]  
[1033]  
[1034]  
[1035]  
[1036]  
[1037]  
[1038]  
[1039]  
[1040]  
[1041]  
[1042]  
[1043]  
[1044]  
[1045]  
[1046]  
[1047]  
[1048]  
[1049]  
[1050]  
[1051]  
[1052]  
[1053]  
[1054]  
[1055]  
[1056]  
[1057]  
[1058]  
[1059]  
[1060]  
[1061]  
[1062]  
[1063]  
[1064]  
[1065]  
[1066]  
[1067]  
[1068]  
[1069]  
[1070]  
[1071]  
[1072]  
[1073]  
[1074]  
[1075]  
[1076]  
[1077]  
[1078]  
[1079]  
[1080]  
[1081]  
[1082]  
[1083]  
[1084]  
[1085]  
[1086]  
[1087]  
[1088]  
[1089]  
[1090]  
[1091]  
[1092]  
[1093]  
[1094]  
[1095]  
[1096]  
[1097]  
[1098]  
[1099]  
[1100]  
[1101]  
[1102]  
[1103]  
[1104]  
[1105]  
[1106]  
[1107]  
[1108]  
[1109]  
[1110]  
[1111]  
[1112]  
[1113]  
[1114]  
[1115]  
[1116]  
[1117]  
[1118]  
[1119]  
[1120]  
[1121]  
[1122]  
[1123]  
[1124]  
[1125]  
[1126]  
[1127]  
[1128]  
[1129]  
[1130]  
[1131]  
[1132]  
[1133]  
[1134]  
[1135]  
[1136]  
[1137]  
[1138]  
[1139]  
[1140]  
[1141]  
[1142]  
[1143]  
[1144]  
[1145]  
[1146]  
[1147]  
[1148]  
[1149]  
[1150]  
[1151]  
[1152]  
[1153]  
[1154]  
[1155]  
[1156]  
[1157]  
[1158]  
[1159]  
[1160]  
[1161]  
[1162]  
[1163]  
[1164]  
[1165]  
[1166]  
[1167]  
[1168]  
[1169]  
[1170]  
[1171]  
[1172]  
[1173]  
[1174]  
[1175]  
[1176]  
[1177]  
[1178]  
[1179]  
[1180]  
[1181]  
[1182]  
[1183]  
[1184]  
[1185]  
[1186]  
[1187]  
[1188]  
[1189]  
[1190]  
[1191]  
[1192]  
[1193]  
[1194]  
[1195]  
[1196]  
[1197]  
[1198]  
[1199]  
[1200]  
[1201]  
[1202]  
[1203]  
[1204]  
[1205]  
[1206]  
[1207]  
[1208]  
[1209]  
[1210]  
[1211]  
[1212]  
[1213]  
[1214]  
[1215]  
[1216]  
[1217]  
[1218]  
[1219]  
[1220]  
[1221]  
[1222]  
[1223]  
[1224]  
[1225]  
[1226]  
[1227]  
[1228]  
[1229]  
[1230]  
[1231]  
[1232]  
[1233]  
[1234]  
[1235]  
[1236]  
[1237]  
[1238]  
[1239]  
[1240]  
[1241]  
[1242]  
[1243]  
[1244]  
[1245]  
[1246]  
[1247]  
[1248]  
[1249]  
[1250]  
[1251]  
[1252]  
[1253]  
[1254]  
[1255]  
[1256]  
[1257]  
[1258]  
[1259]  
[1260]  
[1261]  
[1262]  
[1263]  
[1264]  
[1265]  
[1266]  
[1267]  
[1268]  
[1269]  
[1270]  
[1271]  
[1272]  
[1273]  
[1274]  
[1275]  
[1276]  
[1277]  
[1278]  
[1279]  
[1280]  
[1281]  
[1282]  
[1283]  
[1284]  
[1285]  
[1286]  
[1287]  
[1288]  
[1289]  
[1290]  
[1291]  
[1292]  
[1293]  
[1294]  
[1295]  
[1296]  
[1297]  
[1298]  
[1299]  
[1300]  
[1301]  
[1302]  
[1303]  
[1304]  
[1305]  
[1306]  
[1307]  
[1308]  
[1309]  
[1310]  
[1311]  
[1312]  
[1313]  
[1314]  
[1315]  
[1316]  
[1317]  
[1318]  
[1319]  
[1320]  
[1321]  
[1322]  
[1323]  
[1324]  
[1325]  
[1326]  
[1327]  
[1328]  
[1329]  
[1330]  
[1331]  
[1332]  
[1333]  
[1334]  
[1335]  
[1336]  
[1337]  
[1338]  
[1339]  
[1340]  
[1341]  
[1342]  
[1343]  
[1344]  
[1345]  
[1346]  
[1347]  
[1348]  
[1349]  
[1350]  
[1351]  
[1352]  
[1353]  
[1354]  
[1355]  
[1356]  
[1357]  
[1358]  
[1359]  
[1360]  
[1361]  
[1362]  
[1363]  
[1364]  
[1365]  
[1366]  
[1367]  
[1368]  
[1369]  
[1370]  
[1371]  
[1372]  
[1373]  
[1374]  
[1375]  
[1376]  
[1377]  
[1378]  
[1379]  
[1380]  
[1381]  
[1382]  
[1383]  
[1384]  
[1385]  
[1386]  
[1387]  
[1388]  
[1389]  
[1390]  
[1391]  
[1392]  
[1393]  
[1394]  
[1395]  
[1396]  
[1397]  
[1398]  
[1399]  
[1400]  
[1401]  
[1402]  
[1403]  
[1404]  
[1405]  
[1406]  
[1407]  
[1408]  
[1409]  
[1410]  
[1411]  
[1412]  
[1413]  
[1414]  
[1415]  
[1416]  
[1417]  
[1418]  
[1419]  
[1420]  
[1421]  
[1422]  
[1423]  
[1424]  
[1425]  
[1426]  
[1427]  
[1428]  
[1429]  
[1430]  
[1431]  
[1432]  
[1433]  
[1434]  
[1435]  
[1436]  
[1437]  
[1438]  
[1439]  
[1440]  
[1441]  
[1442]  
[1443]  
[1444]  
[1445]  
[1446]  
[1447]  
[1448]  
[1449]  
[1450]  
[1451]  
[1452]  
[1453]  
[1454]  
[1455]  
[1456]  
[1457]  
[1458]  
[1459]  
[1460]  
[1461]  
[1462]  
[1463]  
[1464]  
[1465]  
[1466]  
[1467]  
[1468]  
[1469]  
[1470]  
[1471]  
[1472]  
[1473]  
[1474]  
[1475]  
[1476]  
[1477]  
[1478]  
[1479]  
[1480]  
[1481]  
[1482]  
[1483]  
[1484]  
[1485]  
[1486]  
[1487]  
[1488]  
[1489]  
[1490]  
[1491]  
[1492]  
[1493]  
[1494]  
[1495]  
[1496]  
[1497]  
[1498]  
[1499]  
[1500]  
[1501]  
[1502]  
[1503]  
[1504]  
[1505]  
[1506]  
[1507]  
[1508]  
[1509]  
[1510]  
[1511]  
[1512]  
[1513]  
[1514]  
[1515]  
[1516]  
[1517]  
[1518]  
[1519]  
[1520]  
[1521]  
[1522]  
[1523]  
[1524]  
[1525]  
[1526]  
[1527]  
[1528]  
[1529]  
[1530]  
[1531]  
[1532]  
[1533]  
[1534]  
[1535]  
[1536]  
[1537]  
[1538]  
[1539]  
[1540]  
[1541]  
[1542]  
[1543]  
[1544]  
[1545]  
[1546]  
[1547]  
[1548]  
[1549]  
[1550]  
[1551]  
[1552]  
[1553]  
[1554]  
[1555]  
[1556]  
[1557]  
[1558]  
[1559]  
[1560]  
[1561]  
[1562]  
[1563]  
[1564]  
[1565]  
[1566]  
[1567]  
[1568]  
[1569]  
[1570]  
[1571]  
[1572]  
[1573]  
[1574]  
[1575]  
[1576]  
[1577]  
[1578]  
[1579]  
[1580]  
[1581]  
[1582]  
[1583]  
[1584]  
[1585]  
[1586]  
[1587]  
[1588]  
[1589]  
[1590]  
[1591]  
[1592]  
[1593]  
[1594]  
[1595]  
[1596]  
[1597]  
[1598]  
[1599]  
[1600]  
[1601]  
[1602]  
[1603]  
[160

ID Q95VB7 PRELIMINARY; PRT; 608 AA.  
AC Q95VB7;  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Albumin.  
OS Schistosoma mansoni (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
OC Schistosomatidae; Schistosomatidae; Schistosoma.  
OX NCBI\_TaxID=6183;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Osman A., Asahi H., Staderker M.J., Loverde P.T.;  
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg  
RT component in murine infection with Schistosoma mansoni";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF418550; AAL08579.1;  
DR GO; GO:0005655; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR002486; Serum albumin; 1.  
DR PRODOM; PD00103; ALBUMIN; 3.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 608 AA; 68225 MW; 55EABE28E1C66E54 CRC64;  
Query Match 100.0%; Score 38; DB 5; Length 608;  
Best Local Similarity 100.0%; Pred. No. 9.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VADESAEN 8  
DB 78 VADESAEN 85  
RESULT 3  
ID P94805 PRELIMINARY; PRT; 157 AA.  
AC P94805;  
DT 01-MAY-1997 (TRENBLrel. 03, Created)  
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Haloflex sp. (strain Aa 2.2).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Haloferax.  
OX NCBI\_TaxID=2254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-SB1;  
RX MEDLINE=20223653; PubMed=10760168;  
RA Holmes M.L., Dyal-Smith M.L.;  
RT "Sequence and expression of a halobacterial beta-galactosidase gene";  
RL Mol. Microbiol. 36:114-122 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN-SB1;  
RA Holmes M.L., Dyal-Smith M.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U70664; AAB40124.1; -;  
DR PIR; T44794; T44794.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR008957; FN III-like.  
DR Pfam; PF00041; fn3; 1.  
DR SMART; SM00060; FN3; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 157 AA; 17392 MW; 65445110F57407FD CRC64;  
Query Match 84.2%; Score 32; DB 1; Length 157;  
Best Local Similarity 87.5%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 102 VADESAEN 109  
RESULT 4  
ID Q8CG74 PRELIMINARY; PRT; 205 AA.  
AC Q8CG74;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Albumin (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=129/SvevTACfBr;  
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ277794; CAC81303.1; -;  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 1.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 1.  
FT NON TER 205 205  
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;  
Query Match 84.2%; Score 32; DB 1; Length 205;  
Best Local Similarity 87.5%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VADESAEN 8  
DB 78 VADESAEN 85  
RESULT 5  
ID Q9PKX9 PRELIMINARY; PRT; 253 AA.  
AC Q9PKX9;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE RNA polymerase sigma factor.  
GN TC0331.  
OS Chlamydia muridarum.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=MoPn/Ni99;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
RA Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
RT pneumoniae RA39";  
RT Nucleic Acids Res. 28:1397-1406 (2000).  
CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES  
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND  
CC THEN IS RELEASED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.  
DR EMBL; AE002301; AAF39194.1; -;  
DR PIR; G81713; G81713.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.08772 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61  
Perfect score: 38  
Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 38    | 100.0       | 396    | 4  | Q81UK7      |
| 2          | 38    | 100.0       | 608    | 5  | Q95VB7      |
| 3          | 32    | 84.2        | 157    | 1  | P94805      |
| 4          | 32    | 84.2        | 205    | 11 | Q8CG74      |
| 5          | 32    | 84.2        | 253    | 16 | Q9PKX9      |
| 6          | 32    | 84.2        | 305    | 10 | Q932D1      |
| 7          | 32    | 84.2        | 385    | 10 | Q9LFB0      |
| 8          | 32    | 84.2        | 576    | 11 | Q8C7C7      |
| 9          | 32    | 84.2        | 608    | 11 | Q8C7H3      |
| 10         | 32    | 84.2        | 649    | 5  | Q8WQC7      |
| 11         | 32    | 84.2        | 1128   | 5  | O62235      |
| 12         | 32    | 84.2        | 1702   | 5  | O8WQC8      |
| 13         | 31    | 81.6        | 75     | 2  | Q9F428      |
| 14         | 31    | 81.6        | 75     | 16 | Q8Y784      |
| 15         | 31    | 81.6        | 215    | 16 | Q8EDP3      |
| 16         | 31    | 81.6        | 232    | 10 | Q9M0M3      |

|    |    |      |      |    |        |
|----|----|------|------|----|--------|
| 17 | 31 | 81.6 | 236  | 16 | Q8E684 |
| 18 | 31 | 81.6 | 236  | 16 | Q8E0L1 |
| 19 | 31 | 81.6 | 274  | 16 | Q7VST1 |
| 20 | 31 | 81.6 | 282  | 16 | Q86U2  |
| 21 | 31 | 81.6 | 289  | 2  | O8857  |
| 22 | 31 | 81.6 | 785  | 10 | O64756 |
| 23 | 31 | 81.6 | 918  | 10 | Q8SA86 |
| 24 | 31 | 81.6 | 954  | 10 | O23575 |
| 25 | 31 | 81.6 | 1058 | 10 | Q8LP18 |
| 26 | 31 | 81.6 | 1657 | 5  | Q9XUC3 |
| 27 | 30 | 78.9 | 107  | 5  | Q9VU1  |
| 28 | 30 | 78.9 | 109  | 2  | Q04818 |
| 29 | 30 | 78.9 | 189  | 10 | Q93WN1 |
| 30 | 30 | 78.9 | 219  | 2  | Q9EUE8 |
| 31 | 30 | 78.9 | 221  | 10 | O82354 |
| 32 | 30 | 78.9 | 232  | 12 | Q8QS76 |
| 33 | 30 | 78.9 | 265  | 17 | Q9HS56 |
| 34 | 30 | 78.9 | 301  | 2  | Q46775 |
| 35 | 30 | 78.9 | 308  | 2  | P71522 |
| 36 | 30 | 78.9 | 309  | 10 | Q9LQY6 |
| 37 | 30 | 78.9 | 325  | 2  | Q9S143 |
| 38 | 30 | 78.9 | 329  | 16 | Q93GR9 |
| 39 | 30 | 78.9 | 339  | 2  | P72427 |
| 40 | 30 | 78.9 | 345  | 16 | Q7VL69 |
| 41 | 30 | 78.9 | 387  | 10 | Q9SHD8 |
| 42 | 30 | 78.9 | 451  | 5  | Q86SK4 |
| 43 | 30 | 78.9 | 453  | 5  | Q9WI40 |
| 44 | 30 | 78.9 | 466  | 5  | Q93712 |
| 45 | 30 | 78.9 | 475  | 10 | Q9SKT8 |

## ALIGNMENTS

RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.

ID Q81UK7: (TREMBL:23, Created)  
AC Q81UK7: (TREMBL:23, Last sequence update)  
DT 01-MAR-2003 (TREMBL:23, Last sequence update)  
DT 01-OCT-2003 (TREMBL:25, Last annotation update)  
DE Similar to serum albumin precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035969; AAH35969.1; -  
DR GO; GO:0005635; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_prot; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 38; DB 4; Length 396;  
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

QY 1 VADESAEN 8

DB 78 VADESAEN 85

RESULT 2

Q95VB7



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; M25787; AAA53283.1; -;  
 DR EMBL; M33129; AAA30224.1; -;  
 DR EMBL; X52584; CAA36814.1; -;  
 DR PIR; A30229; A30229.  
 DR InterPro; IPR008882; Trypano PARP.  
 DR Pfam; PF05887; Trypano PARP; 1.  
 KW Signal; Antigen; Repeat; GPI-anchor; Lipoprotein.  
 FT SIGNAL 1 27  
 FT CHAIN 28 92 PROCYCLIC FORM SPECIFIC POLYPEPTIDE A-  
 FT ALPHA.  
 FT PROPEP 93 114  
 FT DOMAIN 48 85 6 X 5 AA TANDEM REPEATS OF G-P-E-T.  
 FT LIPID 92 92 GPI-anchor amidated glycine.  
 SQ SEQUENCE 114 AA; 11611 MW; FFF2690DAAAE445E CRC64;

Query Match 76.3%; Score 29; DB 1; Length 114;  
 Best Local Similarity 85.7%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ADESAEN 8  
 DB 28 ADESAEN 34

Search completed: April 19, 2004, 11:52:46  
 Job time : 1.58356 secs

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M23391; AAA28487.1; -  
CC PIR; A30817; A30817.  
CC TRANSFAC; T00768; -  
CC FlyBase; FBgn005642; wdn.  
CC InterPro; IPR007087; Znf\_C2H2.  
CC Pfam; PF00096; zf-C2H2; 7.  
CC ProDom; PD000003; Znf\_C2H2; 1.  
CC SMART; SM00355; Znf\_C2H2; 7.  
CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 6.  
CC PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 6.  
CC Developmental protein; Zinc-finger; Metal-binding; DNA-binding;  
KW Transcription regulation; Repeat; Nuclear protein.  
KW SER/THR-RICH.  
FT DOMAIN 52 85  
FT DOMAIN 115 162  
FT DOMAIN 141 213  
FT DOMAIN 169 186  
FT ZN\_FING 268 292  
FT ZN\_FING 298 320  
FT ZN\_FING 330 351  
FT ZN\_FING 357 379  
FT ZN\_FING 385 407  
FT ZN\_FING 413 435  
FT ZN\_FING 441 463  
FT ZN\_FING 469 492  
FT DOMAIN 521 531  
FT DOMAIN 593 669  
FT DOMAIN 632 647  
FT DOMAIN 669 677  
SQ SEQUENCE 868 AA; 95360 MW; 0234DD17F02BF20 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 868;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VADESAEN 8  
Db 90 VLDESAQN 97

RESULT 14  
ACP\_BUCBP  
ID\_ACP\_BUCBP STANDARD; PRT; 78 AA.  
AC P59449;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Acyl carrier protein (ACP).  
GN ACP OR BAP322.  
OS Buchnera aphidicola (subsp. Baizhongia pistaciae).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Buchnera.  
OX NCBI\_TaxID=135842;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426901; PubMed=12522265;  
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,  
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,  
RA Tanames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;  
RT "Reductive genome evolution in Buchnera aphidicola.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
CC -!- FUNCTION: Carrier of the growing fatty acid chain in fatty acid  
CC biosynthesis (By similarity).  
CC -!- PATHWAY: De novo fatty acid biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- PTM: 4'-phosphopantetheine is transferred from CoA to a specific  
CC serine of apo-ACP by acps. This modification is essential for  
CC activity because fatty acids are bound in thioester linkage to the  
CC sulphydryl of the prosthetic group (By similarity).

CC -!- SIMILARITY: Contains 1 acyl carrier domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC HAMAP; HAMAP; MF 01217; -  
CC InterPro; IPR001231; Acyl carrier.  
CC InterPro; IPR006163; Pp\_bind.  
CC InterPro; IPR006162; Ppantne\_S.  
CC Pfam; PF00550; Pp-binding; 1.  
CC TIGR; TIGR; TIGR00517; acyl carrier; 1.  
CC PROSITE; PS00075; ACP\_DOMAIN; 1.  
CC PROSITE; PS00112; PHOSPHOPANTETHEINE; 1.  
CC Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine;  
KW Complete proteome.  
FT BINDING 37 37 PHOSPHOPANTETHEINE (BY SIMILARITY).  
SQ SEQUENCE 78 AA; 9031 MW; CAD574DE83690FD0 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 78;  
Best Local Similarity 62.5%; Pred. No. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VADESAEN 8  
Db 55 ITDESAEN 62

RESULT 15  
PARA\_TRYBB  
ID\_PARA\_TRYBB STANDARD; PRT; 114 AA.  
AC P18764;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Procyclic form specific polypeptide A-alpha precursor (Procyclin A-  
DE alpha) (PARP A-alpha).  
GN PARP A-ALPHA.  
OS Trypanosoma brucei brucei.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89261740; PubMed=2725502;  
RA Mowatt M.R., Wisdom G.S., Clayton C.E.;  
RT "Variation of tandem repeats in the developmentally regulated  
RT procyclic acidic repetitive proteins of Trypanosoma brucei.";  
RL Mol. Cell. Biol. 9:1332-1335(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90258995; PubMed=2342468;  
RA Clayton C.E., Furi J.P., Itzhaki J.E., Bellofatto V., Sherman D.R.,  
RA Wisdom G.S., Vijayasarathy S., Mowatt M.R.;  
RT "Transcription of the procyclic acidic repetitive protein genes of  
RT Trypanosoma brucei.";  
RL Mol. Cell. Biol. 10:3036-3047(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=427;  
RA Vijayasarathy S., Ernest J., Itzhaki J., Sherman D., Mowatt M.R.,  
RA Michels P.A.M., Clayton C.E.;  
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Major surface antigen of procyclic forms.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- DEVELOPMENTAL STAGE: Expressed only at a certain stage during  
CC differentiation in the insect vector.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AB006197; BAA21765.1; -;  
DR PIR; JC5838; JC5838.  
DR HSSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 609 SERUM ALBUMIN.  
FT DOMAIN 25 206 ALBUMIN 1.  
FT DOMAIN 213 338 ALBUMIN 2.  
FT DOMAIN 405 596 ALBUMIN 3.  
FT METAL 28 28 COPPER.  
FT METAL 78 87 BY SIMILARITY.  
FT DISULFID 100 116 BY SIMILARITY.  
FT DISULFID 115 126 BY SIMILARITY.  
FT DISULFID 149 194 BY SIMILARITY.  
FT DISULFID 193 202 BY SIMILARITY.  
FT DISULFID 225 271 BY SIMILARITY.  
FT DISULFID 270 278 BY SIMILARITY.  
FT DISULFID 290 304 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 341 386 BY SIMILARITY.  
FT DISULFID 385 394 BY SIMILARITY.  
FT DISULFID 417 463 BY SIMILARITY.  
FT DISULFID 462 473 BY SIMILARITY.  
FT DISULFID 486 502 BY SIMILARITY.  
FT DISULFID 501 512 BY SIMILARITY.  
FT DISULFID 539 584 BY SIMILARITY.  
FT DISULFID 583 592 BY SIMILARITY.  
SQ SEQUENCE 609 AA; 9CA5F97F67EF1A48 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DESAEN 8  
| | | | |  
DB 81 DESAEN 86

RESULT 12  
SEC9 YEAST  
ID SEC9 YEAST STANDARD; PRT; 651 AA.  
AC P40357;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein transport protein SEC9.  
GN SEC9 OR HSS7 OR YGR009C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=35042722; PubMed=7954793;  
RA Brennwald P., Kearns B., Champion K., Keraenen S.,  
RA Bankaitis V., Novick P.;  
RT "Sec9 is a SNAP-25-like component of a yeast SNARE complex that may  
be the effector of Sec4 function in exocytosis."

RL Cell 79:245-258(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hebling U., Hofmann B., Delli H.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: COMPONENT OF A SNARE COMPLEX THAT MAY BE THE EFFECTOR OF  
CC SEC4 FUNCTION IN EXOCYTOSIS.  
CC -!- SIMILARITY: Belongs to the SNAP-25 family.  
CC -!- SIMILARITY: TO YEAST YMR017W.  
CC -!- SIMILARITY: Contains 2 t-SNARE coiled-coil homology domains.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L34336; AAA35034.1; -;  
DR EMBL; Z72794; CAA96992.1; -;  
DR PIR; A55100; A55100.  
DR GerMOnline; 141321; -;  
DR SGD; S0003241; SEC9.  
DR InterPro; IPR000727; T-SNARE.  
DR SMART; SM00397; t-SNARE; 2.  
DR PROSITE; PS0192; T-SNARE; 2.  
KW Protein transport; Repeat; Coiled coil.  
FT DOMAIN 434 496 T-SNARE COILED-COIL HOMOMOLOGY 1.  
FT DOMAIN 588 650 T-SNARE COILED-COIL HOMOMOLOGY 2.  
SQ SEQUENCE 651 AA; 73623 MW; EA314D73D20A10C7 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 651;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8  
| | | | |  
DB 335 ADKSAEN 341

RESULT 13  
SYCY DROME  
ID SYCY DROME STANDARD; PRT; 868 AA.  
AC P15619;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serendipity locus protein H-1 (Wings-down protein).  
GN SYCY-C OR WDN.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89039875; PubMed=3141791;  
RA Vincent A., Kejzlarova-Lepesant J., Segalat L., Vanicostas C.,  
RA Lepesant J.-A.;  
RT "sry h-1, a new Drosophila melanogaster multifingered protein gene  
showing maternal and zygotic expression."

RL Mol. Cell. Biol. 8:4459-4468(1988).  
CC -!- FUNCTION: THIS MATERNAL AND ZYGOTIC FINGER PROTEIN MAY BELONG TO  
CC A COMPLEX SET OF MULTIFINGERED PROTEINS WHICH PLAY AN IMPORTANT  
CC ROLE IN GENE ACTIVATION OR REGULATION AT EARLY EMBRYONIC STAGES  
CC THROUGH A MAXIMAL ACCUMULATION OF THEIR TRANSCRIPTS (OR PROTEIN  
CC PRODUCT) IN THE MATURE OOCYTE.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DEVELOPMENTAL STAGE: Blastoderm specific.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=97382443; PubMed=9240447;
RX  Zeng Z., Kyaw H., Gakenheimer K.R., Augustus M., Fan P., Zhang X.,
RA  Su K., Carter K.C., Li Y.;
RT  "Cloning, mapping, and tissue distribution of a human homologue of the
RL  mouse jerky gene product.";
RT  Biochem. Biophys. Res. Commun. 236:389-395(1997).
CC  -!- TISSUE SPECIFICITY: Abundantly expressed in the majority of
CC  tissues examined, including brain and skeletal muscle.
CC  -!- SIMILARITY: Contains 1 CENPB domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF004715; AAB65833.1; -.
DR  F1R; JCS594; JCS594.
DR  Genew; HGNC:6200; JRXI.
DR  MIM; 603211; -.
DR  GO; GO:0005634; C:nucleus; TAS.
DR  GO; GO:0007417; P:central nervous system development; TAS.
DR  InterPro; IPR004875; CENP-B.
DR  InterPro; IPR006600; CENPB.
DR  Pfam; PF03184; DDE; 1.
DR  SMART; SMO0674; CENPB; 1.
DR  DOMAIN 104 314 CENP-B.
FT  SEQUENCE 442 AA; 50710 MW; 5EE5430FC54CE488 CRC64;
SQ  SEQUENCE 442 AA; 50710 MW; 5EE5430FC54CE488 CRC64;

Query Match 81.8%; Score 31; DB 1; Length 442;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
Db 375 ADESEN 381
|||||

RESULT 10
SYD_XANAC STANDARD; PRT; 478 AA.
AC Q8PLH2:
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE (HISRS).
GN HISS OR XAC1826.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=20222145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorallo C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,

```

```

RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -!- CATABOLIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; AB011816; AAM36688.1; -.
DR  HAWAP; MF_00127; -.
DR  InterPro; IPR004154; HGTP anticodon.
DR  InterPro; IPR004516; HISS.
DR  InterPro; IPR002314; tRNA-synt 2b.
DR  InterPro; IPR006195; tRNA ligase II.
DR  Pfam; PF03129; HGTP anticodon; 1.
DR  Pfam; PF00587; tRNA-synt 2b; 1.
DR  TIGRFAMs; TIGR00442; HISS; 1.
DR  PROSITE; PS50862; AA tRNA LIGASE II; 1.
DR  Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 478 AA; 52630 MW; B2344B1147990E0E CRC64;

Query Match 78.9%; Score 30; DB 1; Length 478;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ADESAEN 8
Db 79 ADEGAEN 85
|||||

RESULT 11
ALBU_MERUN STANDARD; PRT; 609 AA.
AC O35050:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=981116663; PubMed=9455485;
RA Yoshida K., Seto-Oshihara A., Sinohara H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RL DNA Res. 4:351-354(1997).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----

```

DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport prot.; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00312; ALBUMIN; 3.  
 DR KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 DR SIGNAL 1 18  
 DR FT PROPEP 19 24  
 DR FT CHAIN 25 608  
 DR FT DOMAIN 25 397  
 DR FT DOMAIN 212 397  
 DR FT DOMAIN 404 595  
 DR FT METAL 27 27  
 DR FT DISULFID 77 86  
 DR FT DISULFID 99 115  
 DR FT DISULFID 114 125  
 DR FT DISULFID 148 193  
 DR FT DISULFID 192 201  
 DR FT DISULFID 224 270  
 DR FT DISULFID 269 277  
 DR FT DISULFID 289 303  
 DR FT DISULFID 302 313  
 DR FT DISULFID 340 385  
 DR FT DISULFID 384 393  
 DR FT DISULFID 416 462  
 DR FT DISULFID 461 472  
 DR FT DISULFID 485 501  
 DR FT DISULFID 500 511  
 DR FT DISULFID 538 583  
 DR FT DISULFID 582 591  
 DR SEQUENCE 608 AA; CF5E92647AAFE9A2 CRC64;  
 Query Match 84.2%; Score 32; DB 1; Length 608;  
 Best Local Similarity 87.5%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 DB 78 VADESAEN 85  
 RESULT 8  
 LAG1\_YEAST STANDARD; PRT; 411 AA.  
 AC P38703; 1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Longevity-assurance protein 1 (Longevity assurance factor 1).  
 GN LAG1 OR YHL003C.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=94253121; PubMed=8195187;  
 RA D'Mello N.P., Childress A.M., Franklin D.S., Kale S.P.,  
 RA Pinswasdi C., Jazwinska S.M.;  
 RA "Cloning and characterization of LAG1, a longevity-assurance gene in  
 RT yeast.";  
 RL J. Biol. Chem. 269:15451-15459(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,  
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaaskis E., Vaughan K.,  
 RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,  
 RA Vaubin M.;  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII";  
 RL Science 265:2077-2082(1994).  
 CC -!- FUNCTION: Involved in the aging process. Deletion of LAG1 results  
 CC in a pronounced increase (approximately 50%) in mean and in  
 CC maximum life span.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Contains 1 TLC (TRAM/LAG1/CLNS) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U08133; AAA21579.1; -;  
 DR EMBL; U10555; AAB84429.1; -;  
 DR PIR; S46800; S46800.  
 DR Germonline; 139270; -;  
 DR SGD; S0000995; LAG1.  
 DR GO; GO:0005783; C:Endoplasmic reticulum; IDA.  
 DR GO; GO:0007574; P:cell aging (sensu Saccharomycetes); IMP.  
 DR GO; GO:0046513; P:ceramide biosynthesis; IMP.  
 DR InterPro; IPR005547; LAG1.  
 DR InterPro; IPR006634; TLC.  
 DR SMART; PF03798; LAG1; 1.  
 DR SMART; SM00724; TLC; 1.  
 DR PROSITE; PS50922; TLC; 1.  
 KW Transmembrane.  
 FT TRANSMEM 82 102 POTENTIAL.  
 FT TRANSMEM 135 155 POTENTIAL.  
 FT TRANSMEM 177 197 POTENTIAL.  
 FT TRANSMEM 212 232 POTENTIAL.  
 FT TRANSMEM 252 272 POTENTIAL.  
 FT TRANSMEM 297 317 POTENTIAL.  
 FT TRANSMEM 356 376 POTENTIAL.  
 FT DOMAIN 168 384 TLC.  
 FT CONFLICT 173 174 ML -> IV (IN REF. 2).  
 FT CONFLICT 220 220 F -> C (IN REF. 2).  
 FT CONFLICT 301 411 VTFPFVGLVFFVFIYLRHVVNIILSVLTERHGNVYL  
 NPATQQYKWSIPIVFVLIALLQVLNLYWFLILLYRL  
 IWQIQKDSRSDSDSEAESEKCE -> TEISGIWE  
 KQIDSNDNPTALSPNETSKQVXKPLVLLNPTENRAL  
 LEAIKSRVPIAIDTDEPSLVTPYIPGNDLSLVNFL  
 GVILARAGQGLQNLARNNEK (IN REF. 1).  
 SQ SEQUENCE 411 AA; 48454 MW; 91676D58AC053F3C CRC64;  
 Query Match 81.6%; Score 31; DB 1; Length 411;  
 Best Local Similarity 85.7%; Pred. No. 39;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ADESAEN 8  
 DB 397 SDESAEN 403  
 RESULT 9  
 JKRL\_HUMAN  
 ID JKRL\_HUMAN STANDARD; PRT; 442 AA.  
 AC Q9V4A0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Jerky protein homolog like (HHMJG).  
 GN JKRL  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Locuallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Kulyk S.W.,  
 RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahay J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 99-516 FROM N.A.  
 RX MEDLINE=88216123; PubMed=2452956;  
 RA Minghetti P.P., Law S.W., Dugaiczky A.,  
 RA "The rate of molecular evolution of alpha-fetoprotein approaches that  
 RT of pseudogenes.";  
 RL Mol. Biol. Evol. 2:347-358 (1985).  
 RN [5]  
 RP SEQUENCE OF 477-551 FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=90269606; PubMed=1971802;  
 RA Boccaccio C., Deschattre J., Neunier-Rotival M.,  
 RA "Empty and occupied insertion site of the truncated LINE-1 repeat  
 RT located in the mouse serum albumin-encoding gene.";  
 RL Gene 88:181-186 (1990).  
 RN [6]  
 RP SEQUENCE OF 25-44.  
 RC TISSUE=Liver;  
 RX MEDLINE=93162044; PubMed=1286668;  
 RA Glometti C.S., Taylor J., Tollaksen S.L.,  
 RA "Mouse liver protein database: a catalog of proteins detected by two-  
 RT dimensional gel electrophoresis.";  
 RL Electrophoresis 13:970-991 (1992).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AJ011413; CAA09617.1; -;  
 DR EMBL; AK010025; BAB26650.1; -;  
 DR EMBL; BC049971; AAA49971.1; -;  
 DR EMBL; M16111; AAA37190.1; -;  
 DR EMBL; X13060; CAA31458.1; -;  
 DR PIR; A05139; A05139.  
 DR HSP; P02768; 1E7B.  
 DR SWISS-2DFAGE; P07724; MOUSE.  
 DR MGD; MGI:87991; Albl.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SMO0103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Metal-binding; lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT PROPEP 19 24

FT CHAIN 25 608 SERUM ALBUMIN.  
 FT DOMAIN 25 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.  
 FT DOMAIN 404 595 ALBUMIN 3.  
 FT METAL 27 27 COPPER.  
 FT DISULFID 77 86 BY SIMILARITY.  
 FT DISULFID 99 115 BY SIMILARITY.  
 FT DISULFID 114 125 BY SIMILARITY.  
 FT DISULFID 148 193 BY SIMILARITY.  
 FT DISULFID 192 201 BY SIMILARITY.  
 FT DISULFID 224 270 BY SIMILARITY.  
 FT DISULFID 269 277 BY SIMILARITY.  
 FT DISULFID 289 303 BY SIMILARITY.  
 FT DISULFID 302 313 BY SIMILARITY.  
 FT DISULFID 340 385 BY SIMILARITY.  
 FT DISULFID 384 393 BY SIMILARITY.  
 FT DISULFID 416 462 BY SIMILARITY.  
 FT DISULFID 461 472 BY SIMILARITY.  
 FT DISULFID 485 501 BY SIMILARITY.  
 FT DISULFID 500 511 BY SIMILARITY.  
 FT DISULFID 538 583 BY SIMILARITY.  
 FT DISULFID 582 591 BY SIMILARITY.  
 FT CONFLICT 27 27 H -> D (IN REF. 6).  
 FT CONFLICT 33 33 H -> D (IN REF. 6).  
 FT CONFLICT 41 41 Q -> I (IN REF. 6).  
 SQ SEQUENCE 608 AA; 68692 MW; 292F7C7BED3A61B4 CRC64;  
 Query Match 84.2%; Score 32; DB 1; Length 608;  
 Best Local Similarity 87.5%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 DB 78 VADESAEN 85  
 RESULT 7  
 ALBU RABIT  
 ID ALBU RABIT STANDARD; PRT; 608 AA.  
 AC P49065;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Liver;  
 RA Sheffield W.P., Syed S., Schuyler P.D.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U18344; AAB58347.1; -;  
 DR HSP; P02768; 1E7B.

```

SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
Query Match 92.1%; Score 35; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
    |||:|
DB 78 VADENAEN 85

RESULT 5
ALBU HORSE
ID ALBU HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VPB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X74045; CAA52194.1; -
CC F01; S34053; ABH08.
CC F01; S34053; ABH08.
CC HSP; P02768; 1E7B.
CC InterPro: IPR000264; Serum albumin.
CC Pfam: PF00273; transport prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
CC SIGNAL 1 18
CC PROPEP 19 24
CC CHAIN 25 607
CC DOMAIN 25 204
CC DOMAIN 211 396
CC DOMAIN 403 594
CC METAL 27 27
CC METAL 77 86
CC DISULFID 99 115
CC DISULFID 114 125
CC DISULFID 147 192
CC DISULFID 191 200
CC DISULFID 223 269

```

```

FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 89.5%; Score 34; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
    |||:|
DB 79 ADESAEN 85

RESULT 6
ALBU MOUSE
ID ALBU MOUSE STANDARD; PRT; 608 AA.
AC P07724; O61802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALB1 OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA van Reeth T., Cabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayaishizaki Y.
RA "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP Nature 409:685-690(2001).
RC SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=23388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```





RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczkyk A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao P., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,  
RA Basak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RN SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RN SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.Q.;  
RL (in) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RN SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Farrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RN SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=96140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tanaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RN SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylserine sulfonate";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RN SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RN DISULFIDE BONDS.  
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RN BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RN VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RN VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese allolalbumins";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RN VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RN DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115952; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
RT human serum albumin whose precursor has an aberrant signal peptidase  
RT cleavage site";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
RT Italy";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RN VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,

```
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 484 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; E45C871A670E740B CRC64;

Query Match 100.0%; Score 38; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 70 VADESAEN 77

RESULT 2
ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Baldwin G.S., Weinstein J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X12422; CAA30970.1; --
CC EMBL; M36787; AAA30988.1; --
CC PIR; S01382; ABPGS.
CC HSSP; P02768; 1B7H.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
```

```
DR PRINTS; P00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match 100.0%; Score 38; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 76 VADESAEN 83

RESULT 3
ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; Q55574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minihetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=80081982; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.583564 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61  
Perfect score: 38  
Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 38    | 100.0       | 600    | 1     | ALBU MACMU  |
| 2          | 38    | 100.0       | 605    | 1     | ALBU PIG    |
| 3          | 38    | 100.0       | 609    | 1     | ALBU HUMAN  |
| 4          | 35    | 92.1        | 608    | 1     | ALBU RAT    |
| 5          | 34    | 89.5        | 607    | 1     | ALBU HORSE  |
| 6          | 32    | 84.2        | 608    | 1     | ALBU MOUSE  |
| 7          | 32    | 84.2        | 608    | 1     | ALBU RABIT  |
| 8          | 31    | 81.6        | 411    | 1     | LAGI YEAST  |
| 9          | 31    | 81.6        | 442    | 1     | JRKL HUMAN  |
| 10         | 30    | 78.9        | 478    | 1     | SYH XANAC   |
| 11         | 30    | 78.9        | 609    | 1     | ALBU MERUN  |
| 12         | 30    | 78.9        | 651    | 1     | SECS YEAST  |
| 13         | 30    | 78.9        | 868    | 1     | SRYC DROME  |
| 14         | 29    | 76.3        | 78     | 1     | ACP EUCBP   |
| 15         | 29    | 76.3        | 114    | 1     | PARA TRYBS  |
| 16         | 29    | 76.3        | 186    | 1     | KPTA AGRT5  |
| 17         | 29    | 76.3        | 263    | 1     | MAZG ECOLI  |
| 18         | 29    | 76.3        | 311    | 1     | SREL CAEEL  |
| 19         | 29    | 76.3        | 409    | 1     | GATD THEAC  |
| 20         | 29    | 76.3        | 410    | 1     | CGEL HUMAN  |
| 21         | 29    | 76.3        | 425    | 1     | HISX HALN1  |
| 22         | 29    | 76.3        | 490    | 1     | AMPI STRCO  |
| 23         | 29    | 76.3        | 608    | 1     | ALBU FELCA  |
| 24         | 29    | 76.3        | 1148   | 1     | RFCI HUMAN  |
| 25         | 29    | 76.3        | 1406   | 1     | TOPI CANGA  |
| 26         | 29    | 76.3        | 1463   | 1     | PA2R BOVIN  |
| 27         | 29    | 76.3        | 2003   | 1     | YDBA ECOLI  |
| 28         | 28    | 73.7        | 76     | 1     | YNCU ECOLI  |
| 29         | 28    | 73.7        | 148    | 1     | FLAV DESDE  |
| 30         | 28    | 73.7        | 158    | 1     | PR1 ASPOF   |
| 31         | 28    | 73.7        | 179    | 1     | GRPE STRMU  |
| 32         | 28    | 73.7        | 212    | 1     | SN25 DROME  |
| 33         | 28    | 73.7        | 280    | 1     | AROK ARCFU  |

## RESULT 1

| ID | ALBU MACMU   | STANDARD; | PRT; | 600 AA.        |
|----|--|-----------|------|----------------|
| AC | Q28522;  |           |      |                |
| DT | 01-NOV-1997 (Rel. 35, Created)   |           |      |                |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update)  |           |      |                |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update)  |           |      |                |
| DE | Serum albumin precursor (Fragment).  |           |      |                |
| GN | ALB.   |           |      |                |
| OS | Macaca mulatta (Rhesus macaque).   |           |      |                |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |      |                |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  |           |      |                |
| OC | Cercopitheidae; Macaca.  |           |      |                |
| OX | NCBI_TaxID=9544;   |           |      |                |
| RN | [1]  |           |      |                |
| RP | SEQUENCE FROM N.A.   |           |      |                |
| RX | MEDLINE=93211971; PubMed=8460152;  |           |      |                |
| RA | Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,   |           |      |                |
| RA | Dwulet J., Putnam F.W.;  |           |      |                |
| RT | "cDNA and protein sequence of polymorphic macaque albumins that differ   |           |      |                |
| RT | in bilirubin binding";   |           |      |                |
| RL | Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  |           |      |                |
| CC | !- FUNCTION: Serum albumin, the main protein of plasma, has a good   |           |      |                |
| CC | binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  |           |      |                |
| CC | hormones, bilirubin and drugs. Its main function is the regulation   |           |      |                |
| CC | of the colloidal osmotic pressure of blood.  |           |      |                |
| CC | !- SUBCELLULAR LOCATION: Secreted.   |           |      |                |
| CC | !- TISSUE SPECIFICITY: Plasma.   |           |      |                |
| CC | !- SIMILARITY: Belongs to the ALB/APP/VDB family.  |           |      |                |
| CC | !- SIMILARITY: Contains 3 albumin domains.   |           |      |                |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |      |                |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |      |                |
| CC | the European Bioinformatics Institute. There are no restrictions on its  |           |      |                |
| CC | use by non-profit institutions as long as its content is in no way   |           |      |                |
| CC | modified and this statement is not removed. Usage by and for commercial  |           |      |                |
| CC | entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |      |                |
| CC | or send an email to <a href="mailto:license@sb-sib.ch">license@sb-sib.ch</a> ).  |           |      |                |
| DR | EMBL; M90463; AAA36906.1; -  |           |      |                |
| DR | PIR; A47391; A47391.   |           |      |                |
| DR | HSP; P02768; 1E7B.   |           |      |                |
| DR | InterPro; IPR000264; Serum_albumin.  |           |      |                |
| DR | Pfam; PF00273; transport_prot; 3.  |           |      |                |
| DR | PRINTS; PR00802; SERUMALBUMIN.   |           |      |                |
| DR | ProDom; PD002486; Serum_albumin; 1.  |           |      |                |
| DR | SMART; SM00103; ALBUMIN; 3.  |           |      |                |
| DR | PROSITE; PS00212; ALBUMIN; 3.  |           |      |                |
| DR | Metal-binding; Lipid-binding; Repeat; Signal; Copper.  |           |      |                |
| FT | NON_TER  | 1         |      |                |
| FT | SIGNAL   | <1        | 10   | BY SIMILARITY. |
| FT | PROPEP   | 11        | 16   | BY SIMILARITY. |
| FT | CHAIN  | 17        | 600  | SERUM ALBUMIN. |
| FT | DOMAIN   | 17        | 197  | ALBUMIN 1.     |
| FT | DOMAIN   | 204       | 389  | ALBUMIN 2.     |
| FT | DOMAIN   | 396       | 587  | ALBUMIN 3.     |

P29571 methanobact  
Q889m3 pseudomonas  
P52350 human herpes  
P34392 caenorhabdi  
Q830u2 enterococcu  
O59636 pyrococcus  
Q9ptn4 xenopus lae  
P44001 haemophilus  
P30183 arabidopsis  
Q8p288 methanosarc  
P05039 escherichia  
Q92q12 rhizobium m

## ALIGNMENTS

Best Local Similarity 87.5%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 176 VADSAEN 183

QY 2 ADESAEN 8  
DB 375 ADESEN 381

Search completed: April 19, 2004, 12:02:22  
Job time : 2.97507 secs

## RESULT 14

S46800  
LAG1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YHL003c  
C:Species: Saccharomyces cerevisiae  
C>Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 23-Mar-2001  
C:Accession: S46800; A54012  
R:Favella, T.  
submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of S. cerevisiae cosmid 9780.  
A:Reference number: S46797  
A:Accession: S46800  
A:Molecule type: DNA  
A:Residues: 1-411 <FAV>  
A:Cross-references: EMBL:U10555; NID:G500813; PIDN:AA68429.1; PID:G500820; MIPS:YHL003c  
R:D'Amico, N.P.; Childress, A.M.; Franklin, D.S.; Kale, S.P.; Pineswadi, C.; Jazwinski, J. Biol. Chem. 269, 15451-15459, 1994  
A:Title: Cloning and characterization of LAG1, a longevity-assurance gene in yeast.  
A:Reference number: A54012; MUID:94253121; PMID:8195187  
A:Accession: A54012  
A:Molecule type: DNA  
A:Residues: 1-172, 'IV', 175-219, 'C', 221-300, 'TEISGI', 314, 'EKQE', 315, 'DSNDNPTE', 324, 'A', 32  
'A', 381, 'AGQR', 386, 'L', 388, 'NRLARNNEK' <DAM>  
A:Cross-references: GB:U08133  
C:Genetics:  
A:Gene: SGD:LAG1  
A:Cross-references: SGD:S0000995; MIPS:YHL003c  
A:Map position: 8L  
C:Function:  
A:Description: involved in determination of longevity  
C:Superfamily: hypothetical protein YXL008c  
C:Keywords: transmembrane protein

Query Match 81.6%; Score 31; DB 2; Length 411;  
Best Local Similarity 85.7%; Pred. No. 53;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8  
DB 397 SDESAEN 403

## RESULT 15

JC5594  
jerky gene protein homolog - human  
C:Species: Homo sapiens (man)  
C>Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 05-Nov-1999  
C:Accession: JC5594  
R:Zeng, Z.; Kyaw, H.; Gakenheimer, K.R.; Augustus, M.; Fan, P.; Zhang, X.; Su, K.; Carte  
Biochem. Biophys. Res. Commun. 236, 389-395, 1997  
A:Title: Cloning, mapping, and tissue distribution of a human homologue of the mouse jer  
A:Reference number: JC5594; MUID:97382443; PMID:9240447  
A:Accession: JC5594  
A:Molecule type: mRNA  
A:Residues: 1-442 <ZEN>  
A:Cross-references: DBJ:AF004715; NID:G2314828; PIDN:AA65833.1; PID:G2314829  
A>Note: it is uncertain whether Met-1 or Met-33 is the initiator  
C:Comment: This protein functions as a nuclear regulatory protein.  
C:Genetics:  
A:Gene: hhmjg  
A:Map position: 11q21

Query Match 81.6%; Score 31; DB 2; Length 442;  
Best Local Similarity 85.7%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 231 VSDSAQN 238

RESULT 9

T21861

hypothetical protein F36P2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T21861

R:Cottage, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19479

A:Accession: T21861

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1877 <WIL>

A:Cross-references: EMBL:Z81532; PIDN:CAB04326.1; GSPDB:GN00019; CESP:F36P2.3

A:Experimental source: clone F36P2

C:Genetics:

A:Gene: CESP:F36P2.3

A:Map position: 1

A:Introns: 49/3; 86/2; 112/3; 139/2; 235/3; 284/3; 436/3; 507/2; 566/2; 678/2; 1000/3; 1

Query Match 84.2%; Score 32; DB 2; Length 1877;

Best Local Similarity 75.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8

Db 714 VADENAQN 721

RESULT 10

T16871

hypothetical protein T13H2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000

C:Accession: T16871

R:Wu, X.

submitted to the EMBL Data Library, October 1995

A:Description: The sequence of C. elegans cosmid T13H2.

A:Reference number: Z18593

A:Accession: T16871

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2215 <WUX>

A:Cross-references: EMBL:U39653; NID:G1049397; PID:G1049401; PIDN:AAB52495.1; GSPDB:GN00

A:Experimental source: strain Bristol N2; clone T13H2

C:Genetics:

A:Gene: CESP:T13H2.4

A:Map position: X

A:Introns: 112/2; 136/1; 167/1; 196/1; 649/1; 696/1; 757/1; 850/1; 882/1; 985/1; 1046/1;

Query Match 84.2%; Score 32; DB 2; Length 2215;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAP 7

Db 129 VADESAP 135

RESULT 11

AC1251

hypothetical protein lmo1411 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AC1251

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.N.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1251

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAC99489.1; PID:G16410840; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1411

Query Match 81.6%; Score 31; DB 2; Length 75;

Best Local Similarity 75.0%; Pred. No. 8.8;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8

Db 66 VADDSVEN 73

RESULT 12

D71442

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

A:Variety: columbia

C>Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 05-Dec-1998

C:Accession: D71442

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dir

F.; Wedler, H.; Wedler, E.; Wambutt, R.; Weizenecker, T.; Pohl, T.M.; Terry, N.; Gie

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomene

erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ar

C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: D71442

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-232 <BEV>

A:Cross-references: GB:Z97343; NID:G2245073; PID:G2245054; PID:G2245091

C:Genetics:

A:Map position: 4C09-4G3845

Query Match 81.6%; Score 31; DB 2; Length 232;

Best Local Similarity 85.7%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8

Db 117 SDESAEN 123

RESULT 13

T36578

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T36578

R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z21575

A:Accession: T36578

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-282 <OLI>

A:Cross-references: EMBL:AL049826; PIDN:CAB42715.1; GSPDB:GN00070; SCOREDB:SCH24.13C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOREDB:SCH24.13C

Query Match 81.6%; Score 31; DB 2; Length 282;

```

A:Reference number: Z22843
A:Accession: T44794
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-157 <HL>
A:Cross-references: EMBL:U70664; PIDN:AB40124.1
A:Experimental source: strain SBI

Query Match      84.2%; Score 32; DB 2; Length 157;
Best Local Similarity 87.5%; Pred. No. 12;
Matches      7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 VADESAEN 8
      |||||
DB     102 VADETAEN 109

RESULT 7
RNA polymerase sigma factor, sigma-70 family TC0331 [imported] - Chlamydia muridarum (G81713)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: G81713
C:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, R.C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: G81713
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-253 <JET>
A:Cross-references: GB:AE002301; GB:AE002160; MID:g7190372; PIDN:AAF39194.1; PID:g71903
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0331
C:Superfamily: transcription initiation factor sigmaD; transcription initiation factor

Query Match      84.2%; Score 32; DB 2; Length 253;
Best Local Similarity 75.0%; Pred. No. 19;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VADESAEN 8
      |||||
DB     174 IADERAEN 181

RESULT 8
T45966
hypothetical protein F7J8.240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: Z45966
C:Byvan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Le submitted to the Protein Sequence Database, January 2000
A:Reference number: 223018
A:Accession: T45966
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <BEV>
A:Cross-references: EMBL:AL137189
A:Experimental source: cultivar Columbia; BAC clone F7J8
C:Genetics:
A:Map position: 5
A:introns: 72/3; 136/3; 303/3
A:Note: F7J8.240

Query Match      84.2%; Score 32; DB 2; Length 385;
Best Local Similarity 75.0%; Pred. No. 30;
Matches      6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 VADESAEN 8
      |||||

```

A;Molecule type: protein  
A;Residues: 82-105,'K',107-110 <GAL2>  
A;Note: this variant is designated albumin Vibo Valentia  
A;Accession: A38255  
A;Molecule type: protein  
A;Residues: 76-83,'K',85-106 <GAL3>  
A;Note: this variant is designated albumin Torino  
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A;Title: The structural characterization and bilirubin-binding properties of albumin Her  
A;Reference number: S33298; MUID:93292504; PMID:8513793  
A;Accession: S33298  
A;Molecule type: protein  
A;Residues: 255-263,'E',265-281 <MIN1>  
A;Note: this variant is designated albumin Herborn  
R;Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A;Title: Two albumins with identical electrophoretic mobility are produced by differ  
A;Reference number: S21078; MUID:92190239; PMID:1347703  
A;Accession: S21078  
A;Molecule type: protein  
A;Residues: 354-356,'K',358-378 <MIN2>  
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,  
R;He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A;Title: Atomic structure and chemistry of human serum albumin.  
A;Reference number: A46756; MUID:92334427; PMID:1630489  
A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
R;Brown, J.R.; Shockley, P.; Behrens, P.O.  
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
A;Reference number: A94442  
A;Contents: annotation; Three-dimensional structure and disulfide bonds  
R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A;Title: Disulfide bonds in human serum albumin.  
A;Reference number: A90930  
A;Contents: annotation; disulfide bonds  
R;Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A;Reference number: A90299; MUID:78186630; PMID:656055  
A;Contents: annotation; bilirubin-binding site  
R;Peters, T.; Reed, R.G.  
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,  
A;Title: Serum albumin: conformation and active sites.  
A;Reference number: A94408  
A;Contents: annotation; binding sites  
R;Harper, M.E.; Dugaiczky, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
A;Reference number: A90028; MUID:83279982; PMID:6192711  
A;Contents: annotation; gene position  
R;Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
A;Reference number: A46755; MUID:76257808; PMID:955075  
A;Contents: annotation  
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid  
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A;Title: Identification of Lys(199) as the primary binding site for pyridoxal 5'-phospha  
A;Reference number: A56294; MUID:92183881; PMID:1544460  
A;Contents: annotation  
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in F  
atase activity  
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
C;Comment: A large number of variants of human serum albumin have been described.  
C;Genetics  
A;Gene: GDB:ALB  
A;Cross-references: GDB:118990; OMIM:103600  
A;Map position: 4q11-4q13  
C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-609/Product: serum albumin #status experimental <MPT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;166-174/Product: Kinetensin #status experimental <KIP>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,  
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental  
Query Match 100.0%; Score 38; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. NO. 2.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0;  
QY 1 VADESAEN 8  
DB 78 VADESAEN 85  
RESULT 4  
ABRTS  
serum albumin precursor - rat  
N;Alternate names: preproalbumin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999  
C;Accession: A93872; A92211; A91940; C45800; I57621; A03233  
R;Sargent, T.D.; Yang, M.; Bonner, J.  
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981  
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
A;Reference number: A93872; MUID:81223722; PMID:7017712  
A;Accession: A93872  
A;Molecule type: mRNA  
A;Residues: 1-608 <SAR>  
A;Cross-references: GB:V01222; GB:J00698; NID:955627; PID:CAA24532.1; PID:955628  
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.  
J. Biol. Chem. 252, 6846-6855, 1977  
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Anal  
A;Reference number: A92211; MUID:77249557; PMID:893447  
A;Note: cleavages during protein maturation  
A;Accession: A92211  
A;Molecule type: protein  
A;Residues: 1-38 <STR>  
R;Isemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleava  
A;Reference number: A91946; MUID:78109429; PMID:564345  
A;Accession: A91946  
A;Molecule type: protein  
A;Residues: 223-288;572-608 <IS2>  
A;Note: 262-Leu was also found  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A;Reference number: A90758; MUID:79001617; PMID:80265  
A;Contents: annotation; copper binding  
R;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid protea  
A;Reference number: A45800; MUID:89341405; PMID:2474609  
A;Accession: C45800  
A;Status: Preliminary  
A;Molecule type: protein  
A;Residues: 166-173 <CAR>  
R;Heard, J.

A:Molecule type: protein  
A:Residues: 23-51,'X','53-54','XXGXV' 146,'E',148,'E',150-151,'XVN',155 <LIM>  
A:Experimental source: dental enamel  
A>Note: albumin and other serum proteins are also found in bone  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-22/Domain: propeptide #status predicted <PRO>  
F:23-605/Product: serum albumin #status predicted <MAT>  
F:27-199/Domain: serum albumin repeat homology <SA1>  
F:218-391/Domain: serum albumin repeat homology <SA2>  
F:410-589/Domain: serum albumin repeat homology <SA3>  
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4  
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 38; DB 1; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
|||||  
Db 76 VADESAEN 83

RESULT 3  
ABHUS  
serum albumin precursor [validated] - human  
N:Alternate names: preproalbumin  
N:Contains: kinetensin  
C:Species: Homo sapiens (man)  
C:Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text change 17-Mar-2000  
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36  
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli  
A:Reference number: A93743; MUID:82081882; PMID:6171778  
A:Accession: A93743  
A:Molecule type: mRNA  
A:Residues: 1-419,'K',421-609 <LAW>  
A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23  
R:Dugaiczky, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A:Reference number: A93936; MUID:82105994; PMID:6275391  
A:Accession: A93936  
A:Molecule type: mRNA  
A:Residues: 1-120,'G',122-609 <DUG>  
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
R:Uranio, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A:Reference number: I39427; MUID:86140095; PMID:2419329  
A:Accession: I39427  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <URA>  
A:Cross-references: GB:M13075; NID:G178330; PIDN:AAAS1688.1; PID:G553173  
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A:Reference number: I59286; MUID:94181575; PMID:8134387  
A:Accession: I59286  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 282-290,'KSRPDLQ' <WAT>  
A:Cross-references: GB:G59152; NID:G546032; PIDN:AA30282.1; PID:G546033  
A>Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 589-590,'ALPRRVKNLLQVKLP' <WAD>  
A:Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232  
A>Note: this frame-shift variant is designated albumin Bazzano; four additional variant  
R:Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: G08292  
A:Accession: G01747  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-120,'G',122-455 <MEN>  
A:Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Ke  
A:Reference number: S55314; MUID:95275251; PMID:7755581  
A:Accession: S55314  
A:Molecule type: protein  
A:Residues: 19-27 <LED>  
R:Meloun, B.; Moravsek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A:Title: Complete amino acid sequence of human serum albumin.  
A:Reference number: A91420; MUID:76187907; PMID:1225573  
A:Accession: A91420  
A:Molecule type: protein  
A:Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',399;  
R:Roehr, U.; Spitteller, G.; Tripier, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from t  
A:Reference number: S06422  
A>Note: this paper is in German, with an English abstract  
A:Accession: S06422  
A:Molecule type: protein  
A:Residues: 25-48 <ROE>  
R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A:Title: Mass spectrometric identification of modifications to human serum albumin tre  
A:Reference number: S36882; MUID:93384321; PMID:8373158  
A:Accession: S36882  
A:Molecule type: protein  
A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>  
R:Kausler, E.; Spitteller, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelme  
A:Reference number: S17599; MUID:92126241; PMID:1772598  
A:Accession: S17599  
A:Molecule type: protein  
A:Residues: 25-54;354-357;431-447 <KAU>  
A>Note: 49-Leu was also found  
R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid protea  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: A45800  
A:Molecule type: protein  
A:Residues: 166-173,'L' <CAR>  
R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Y  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-t  
A:Reference number: A03239; MUID:86242180; PMID:3087352  
A:Accession: A03239  
A:Molecule type: protein  
A:Residues: 166-173,'L' <MOG>  
R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins,  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A:Title: Mutations in genetic variants of human serum albumin found in Italy.  
A:Reference number: A38255; MUID:91062352; PMID:2247440  
A:Accession: C38255  
A:Molecule type: protein  
A:Residues: 76-111 <GAL1>  
A:Accession: B38255



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.975069 Seconds

(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61

Perfect score: 38

Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1:\*\*

2: Pir2:\*\*

3: Pir3:\*\*

4: Pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 38    | 100.0       | 600    | 2 A47391 | serum albumin prec |
| 2          | 38    | 100.0       | 603    | 1 ABPQS  | serum albumin prec |
| 3          | 38    | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 4          | 35    | 92.1        | 608    | 1 ABRTS  | serum albumin prec |
| 5          | 34    | 89.5        | 607    | 1 ABHOS  | serum albumin prec |
| 6          | 32    | 84.2        | 157    | 2 T44794 | hypothetical prote |
| 7          | 32    | 84.2        | 253    | 2 G81713 | RNA polymerase sig |
| 8          | 32    | 84.2        | 385    | 2 T45966 | hypothetical prote |
| 9          | 32    | 84.2        | 1877   | 2 T21861 | hypothetical prote |
| 10         | 32    | 84.2        | 2215   | 2 T16871 | hypothetical prote |
| 11         | 31    | 81.6        | 75     | 2 AC1251 | hypothetical prote |
| 12         | 31    | 81.6        | 232    | 2 D71442 | hypothetical prote |
| 13         | 31    | 81.6        | 282    | 2 T36578 | probable membrane  |
| 14         | 31    | 81.6        | 411    | 2 S46800 | LAG1 protein - yea |
| 15         | 31    | 81.6        | 442    | 2 JC5594 | terky gene protein |
| 16         | 31    | 81.6        | 785    | 2 T00474 | hypothetical prote |
| 17         | 31    | 81.6        | 1657   | 2 T25421 | hypothetical prote |
| 18         | 30    | 78.9        | 109    | 2 S32886 | hypothetical prote |
| 19         | 30    | 78.9        | 221    | 2 C84839 | hypothetical prote |
| 20         | 30    | 78.9        | 265    | 2 D84138 | endonuclease III   |
| 21         | 30    | 78.9        | 309    | 2 F86393 | protein T24P13.4   |
| 22         | 30    | 78.9        | 387    | 2 D84885 | hypothetical prote |
| 23         | 30    | 78.9        | 465    | 2 T22141 | hypothetical prote |
| 24         | 30    | 78.9        | 475    | 2 D84533 | hypothetical prote |
| 25         | 30    | 78.9        | 496    | 2 AD1211 | B. subtilis yuek p |
| 26         | 30    | 78.9        | 496    | 2 AD1567 | conserved hypothet |
| 27         | 30    | 78.9        | 508    | 2 T22954 | hypothetical prote |
| 28         | 30    | 78.9        | 609    | 2 JC5838 | albumin - Mongolia |
| 29         | 30    | 78.9        | 651    | 2 A55100 | SEC9 protein - yea |

serendipity (sry h  
procyelic acidic r  
conserved hypothet  
probable RNA 2'-ph  
transcription regu  
mazG protein - Esc  
hypothetical prote  
hypothetical prote  
protein B0495.1 [i  
probable oxygen-in  
cyclin E - human  
histidinol dehydro  
probable membrane  
hypothetical prote  
X-Pro aminopeptida  
LlDBP protein - hu

#### ALIGNMENTS

##### RESULT 1

A47391  
serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwyer, J.; Putnam, I  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bili  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote  
A:Accession: A47391  
A>Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M30463; NID:G342294; PIDN:AAA36906.1; PID:G342295  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:21-194/Domain: serum albumin repeat homology <SA1>  
F:213-386/Domain: serum albumin repeat homology <SA2>  
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 38; DB 2; Length 600;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8

DB 70 VADESAEN 77

##### RESULT 2

ABPQS  
serum albumin precursor - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S01382; A61006  
R:Reinstock, J.; Baldwin, G.S.  
Nucleic Acids Res 16, 9045, 1988  
A:Title: Nucleotide sequence of porcine liver albumin.  
A:Reference number: S01382; MUID:89016562; PMID:3174440  
A:Accession: S01382  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-605 <WEI>  
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798  
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of miner  
A:Reference number: A61006; MUID:89269769; PMID:2728927  
A:Accession: A61006

Db 307 VHTCCCHGDLLECADRADLAKVICENQDSISSKLECKECPLEKSHCIAEVENDEMPA 366  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYIYARRHPDYSVVLLLRILAKTYETTLEKC 360  
Db 367 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYIYARRHPDYSVVLLLRILAKTYETTLEKC 426  
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
Db 427 CAAADPHECYAKVDFEFKPLVEEPQNL 453

## RESULT 15

US-10-153-064-105  
; Sequence 105, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 105  
; LENGTH: 652  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-064-105

Query Match 100.0%; Score 2068; DB 4; Length 652;  
Best Local Similarity 100.0%; Pred. No. 2.1e-203; Indels 0; Gaps 0;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHKSEVAHRFKDIDGENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAB 60  
Db 67 DAHKSEVAHRFKDIDGENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAB 126  
Qy 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPV 120  
Db 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPV 186  
Qy 121 DVMCATPHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCOAAADKAAACLLP 180  
Db 187 DVMCATPHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCOAAADKAAACLLP 246  
Qy 181 KLDELDRDEGKASSAKQBLKCSLQKQGERAFKAWAVARLSQRPKAEFVSKLVTDLT 240  
Db 247 KLDELDRDEGKASSAKQBLKCSLQKQGERAFKAWAVARLSQRPKAEFVSKLVTDLT 306  
Qy 241 VHTCCCHGDLLECADRADLAKVICENQDSISSKLECKECPLEKSHCIAEVENDEMPA 300  
Db 307 VHTCCCHGDLLECADRADLAKVICENQDSISSKLECKECPLEKSHCIAEVENDEMPA 366  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYIYARRHPDYSVVLLLRILAKTYETTLEKC 360  
Db 367 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYIYARRHPDYSVVLLLRILAKTYETTLEKC 426  
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
Db 427 CAAADPHECYAKVDFEFKPLVEEPQNL 453

Search completed: April 19, 2004, 12:05:18  
Job time : 60.3186 secs

```

; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-133

Query Match      100.0%; Score 2068; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 2e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 366
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTLEKC 360
DB 367 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTLEKC 426
QY 361 CAAADPHECVAKVDFEFKPLVEBPQNL 387
DB 427 CAAADPHECVAKVDFEFKPLVEBPQNL 453

RESULT 14
US-10-153-064-99
; Sequence 99, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 99
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-99

Query Match      100.0%; Score 2068; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.1e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

; Sequence 96, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-96

Query Match      100.0%; Score 2068; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.1e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFKAFKFAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

```

;; PRIOR APPLICATION DATA: US 08/256,927  
;; APPLICATION NUMBER: US 08/256,927  
;; FILING DATE: 28-JUL-1994  
;; APPLICATION NUMBER: FR 92/01064  
;; FILING DATE: 31-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38,619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-797-689-2

Query Match 100.0%; Score 2068; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.9e-203;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144  
  
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204  
  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
DB 265 VHTCCGDLLECCADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 324  
  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRALKTYETTLK 360  
DB 325 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRALKTYETTLK 384  
  
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
DB 385 CAADPHECYAKVDFEFPKPLVEEPQNL 411

RESULT 11  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. 6686179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guittion, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43

;; CITY: Collegeville  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19426  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: Macintosh  
;; OPERATING SYSTEM: System 7.1  
;; SOFTWARE: Word 5.1 (patentin)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/984,186  
;; FILING DATE: 29-Oct-2001  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/797,689  
;; FILING DATE: 31-JAN-1997  
;; APPLICATION NUMBER: US 08/256,927  
;; FILING DATE: 28-JUL-1994  
;; APPLICATION NUMBER: FR 92/01064  
;; FILING DATE: 31-JAN-1992  
;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38,619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 100.0%; Score 2068; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.9e-203;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144  
  
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204  
  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
DB 265 VHTCCGDLLECCADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 324  
  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRALKTYETTLK 360  
DB 325 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRALKTYETTLK 384  
  
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
DB 385 CAADPHECYAKVDFEFPKPLVEEPQNL 411

RESULT 12  
US-10-153-064-133

```

; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
; US-09-976-594-977

Query Match      100.0%; Score 2068; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCCAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 264
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384
QY 361 CAADPHECYAKVDFEFPKPLVEBPQNL 387
DB 385 CAADPHECYAKVDFEFPKPLVEBPQNL 411

RESULT 9
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRR
; STREET: 1940 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:

```

```

; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-04075-3

Query Match      100.0%; Score 2068; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCCAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 264
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384
QY 361 CAADPHECYAKVDFEFPKPLVEBPQNL 387
DB 385 CAADPHECYAKVDFEFPKPLVEBPQNL 411

RESULT 10
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Pournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435

```

Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264  
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 384  
QY 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387  
DB 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

## RESULT 6

US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 2069; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.9e-203;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264

QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 384  
QY 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387  
DB 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

## RESULT 7

US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PFS56  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 2069; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.9e-203;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264  
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 384  
QY 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387  
DB 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

## RESULT 8

US-09-976-594-977  
; Sequence 977, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-746-2

Query Match      100.0%; Score 2068; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
DB 361 CAAADPHECYAKVFDEKPLVEBPQNL 387

RESULT 4
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
```

```
Query Match      100.0%; Score 2068; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
DB 361 CAAADPHECYAKVFDEKPLVEBPQNL 387

RESULT 5
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RR
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-222-619-3

Query Match      100.0%; Score 2068; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 389..419
; OTHER INFORMATION: /note= "Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; OTHER INFORMATION: natural HSA"
;
; US-08-153-799-14
;
; Query Match 100.0%; Score 2068; DB 1; Length 585;
; Best Local Similarity 100.0%; Pred. No. 1.8e-203;
; Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
; DB 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
;
; QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
; DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
;
; QY 121 DVMTAFHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
; DB 121 DVMTAFHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
;
; QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
; DB 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
;
; QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
; DB 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
;
; QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
; DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
;
; QY 361 CAADPHCEYAKVDFEKPPLVEEPQNL 387
; DB 361 CAADPHCEYAKVDFEKPPLVEEPQNL 387
;
; RESULT 2
; US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19408-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-702-572-2
;
; Query Match 100.0%; Score 2068; DB 2; Length 585;
; Best Local Similarity 100.0%; Pred. No. 1.8e-203;
; Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
; DB 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
;
; QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
; DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
;
; QY 121 DVMTAFHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
; DB 121 DVMTAFHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
;
; QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
; DB 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
;
; QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
; DB 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
;
; QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
; DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
;
; QY 361 CAADPHCEYAKVDFEKPPLVEEPQNL 387
; DB 361 CAADPHCEYAKVDFEKPPLVEEPQNL 387
;
; RESULT 3
; US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 59.3186 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_1\_387  
Perfect score: 2068  
Sequence: 1 DAHKSEVAHRFKDLGEENFK.....ECYAKVFDEKPLVEEPQNL 387

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/prodata/2/iaa/PCUTUS COMB.pcp.\*  
6: /cgn2\_6/prodata/2/iaa/backfile1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | *<br>Match | Query Length | DB ID | Description       |
|------------|-------|------------|--------------|-------|-------------------|
| 1          | 2068  | 100.0      | 585          | 1     | US-08-153-799-14  |
| 2          | 2068  | 100.0      | 585          | 2     | US-08-702-572-2   |
| 3          | 2068  | 100.0      | 585          | 3     | US-08-769-746-2   |
| 4          | 2068  | 100.0      | 585          | 4     | US-10-153-064-5   |
| 5          | 2068  | 100.0      | 609          | 1     | US-08-222-619-3   |
| 6          | 2068  | 100.0      | 609          | 4     | US-08-897-956A-2  |
| 7          | 2068  | 100.0      | 609          | 4     | US-10-153-064-7   |
| 8          | 2068  | 100.0      | 609          | 4     | US-09-976-594-977 |
| 9          | 2068  | 100.0      | 609          | 5     | PCT-US95-04075-3  |
| 10         | 2068  | 100.0      | 610          | 2     | US-08-797-689-2   |
| 11         | 2068  | 100.0      | 610          | 4     | US-09-984-186-2   |
| 12         | 2068  | 100.0      | 651          | 4     | US-10-153-064-133 |
| 13         | 2068  | 100.0      | 652          | 4     | US-10-153-064-96  |
| 14         | 2068  | 100.0      | 652          | 4     | US-10-153-064-99  |
| 15         | 2068  | 100.0      | 652          | 4     | US-10-153-064-105 |
| 16         | 2068  | 100.0      | 652          | 4     | US-10-153-064-132 |
| 17         | 2068  | 100.0      | 653          | 4     | US-10-153-064-131 |
| 18         | 2068  | 100.0      | 655          | 4     | US-10-153-064-130 |
| 19         | 2068  | 100.0      | 660          | 4     | US-10-153-064-90  |
| 20         | 2068  | 100.0      | 660          | 4     | US-10-153-064-93  |
| 21         | 2068  | 100.0      | 668          | 4     | US-10-153-064-102 |
| 22         | 2068  | 100.0      | 676          | 4     | US-10-153-064-95  |
| 23         | 2068  | 100.0      | 676          | 4     | US-10-153-064-98  |
| 24         | 2068  | 100.0      | 676          | 4     | US-10-153-064-104 |
| 25         | 2068  | 100.0      | 676          | 4     | US-10-153-064-127 |
| 26         | 2068  | 100.0      | 676          | 4     | US-10-153-064-129 |
| 27         | 2068  | 100.0      | 677          | 4     | US-10-153-064-125 |

|    |        |       |      |   |                   |
|----|--------|-------|------|---|-------------------|
| 28 | 2068   | 100.0 | 680  | 4 | US-10-153-064-123 |
| 29 | 2068   | 100.0 | 684  | 4 | US-10-153-064-92  |
| 30 | 2068   | 100.0 | 682  | 4 | US-10-153-064-101 |
| 31 | 2068   | 100.0 | 783  | 1 | US-08-256-938-2   |
| 32 | 2068   | 100.0 | 787  | 1 | US-08-256-938-4   |
| 33 | 2068   | 100.0 | 787  | 2 | US-08-797-689-16  |
| 34 | 2068   | 100.0 | 787  | 4 | US-09-984-186-16  |
| 35 | 2068   | 100.0 | 978  | 4 | US-08-897-956A-3  |
| 36 | 2068   | 100.0 | 1184 | 4 | US-10-153-064-89  |
| 37 | 2064   | 99.8  | 609  | 1 | US-08-433-037-4   |
| 38 | 2062   | 99.7  | 585  | 1 | US-08-448-196A-3  |
| 39 | 2062   | 99.7  | 585  | 2 | US-08-984-176-1   |
| 40 | 1658.5 | 80.2  | 583  | 1 | US-08-448-196A-5  |
| 41 | 1622.5 | 78.5  | 583  | 1 | US-08-448-196A-4  |
| 42 | 1613.5 | 78.0  | 583  | 1 | US-08-448-196A-6  |
| 43 | 1610   | 77.9  | 584  | 1 | US-08-448-196A-7  |
| 44 | 1568   | 75.8  | 582  | 1 | US-08-134-638-1   |
| 45 | 726.5  | 35.1  | 579  | 1 | US-08-448-196A-8  |

# ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5765883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

|                    |
|--------------------|
| Sequence 123, Appl |
| Sequence 92, Appl  |
| Sequence 101, Appl |
| Sequence 2, Appl   |
| Sequence 4, Appl   |
| Sequence 16, Appl  |
| Sequence 3, Appl   |
| Sequence 89, Appl  |
| Sequence 4, Appl   |
| Sequence 3, Appl   |
| Sequence 1, Appl   |
| Sequence 5, Appl   |
| Sequence 6, Appl   |
| Sequence 7, Appl   |
| Sequence 1, Appl   |
| Sequence 8, Appl   |

Search completed: April 19, 2004, 12:54:58  
Job time : 5.89474 secs

; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF542  
; CURRENT APPLICATION NUMBER: US/09/832,501  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-832-501-18

Query Match 100.0%; Score 70; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
DB 76 TVATLRETYGEMAD 89

## RESULT 11

US-09-833-118-18  
; Sequence 18, Application US/09833118  
; Publication No. US20030219875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF544  
; CURRENT APPLICATION NUMBER: US/09/833,118  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-118-18

Query Match 100.0%; Score 70; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
DB 76 TVATLRETYGEMAD 89

## RESULT 12

US-09-833-245-18  
; Sequence 18, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-245-18

Query Match 100.0%; Score 70; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
DB 76 TVATLRETYGEMAD 89

## RESULT 13

US-10-424-999-11  
; Sequence 11, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbitt, Mark  
; APPLICANT: Cameron, Beatrice  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Fusion protein human abrogen  
US-10-424-999-11

Query Match 100.0%; Score 70; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
DB 76 TVATLRETYGEMAD 89

## RESULT 14

US-10-425-000-31  
; Sequence 31, Application US/10425000  
; Publication No. US20040052777A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbitt, Mark  
; APPLICANT: Cameron, Beatrice  
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit  
; TITLE OF INVENTION: Angiogenesis  
; FILE REFERENCE: ST01027-B  
; CURRENT APPLICATION NUMBER: US/10/425,000  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 105



```
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-27

Query Match          100.0%; Score 70; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 3
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08193-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-28

Query Match          100.0%; Score 70; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 4
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match          100.0%; Score 70; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 5
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT: DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match          100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 6
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 5.89474 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_76\_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 70    | 100.0       | 195    | 13    | US-10-074-956-24  |
| 2          | 70    | 100.0       | 241    | 13    | Sequence 24, Appl |
| 3          | 70    | 100.0       | 268    | 13    | Sequence 27, Appl |
| 4          | 70    | 100.0       | 585    | 9     | US-10-074-956-28  |
| 5          | 70    | 100.0       | 585    | 10    | Sequence 28, Appl |
| 6          | 70    | 100.0       | 585    | 10    | Sequence 2, Appl  |
| 7          | 70    | 100.0       | 585    | 10    | Sequence 445, App |
| 8          | 70    | 100.0       | 585    | 10    | Sequence 18, Appl |
| 9          | 70    | 100.0       | 585    | 10    | Sequence 18, Appl |
| 10         | 70    | 100.0       | 585    | 10    | Sequence 445, App |
| 11         | 70    | 100.0       | 585    | 11    | US-09-832-501-18  |
| 12         | 70    | 100.0       | 585    | 11    | Sequence 18, Appl |
| 13         | 70    | 100.0       | 585    | 12    | Sequence 18, Appl |
| 14         | 70    | 100.0       | 585    | 12    | Sequence 11, Appl |
| 15         | 70    | 100.0       | 585    | 12    | Sequence 31, Appl |
|            |       |             |        |       | Sequence 34, Appl |

```

16 70 100.0 585 13 US-10-153-064-5 Sequence 5, Appli
17 70 100.0 585 14 US-10-153-604A-5 Sequence 5, Appli
18 70 100.0 585 14 US-10-319-263-1 Sequence 1, Appli
19 70 100.0 585 14 US-10-319-263-2 Sequence 2, Appli
20 70 100.0 585 14 US-10-414-469-1 Sequence 1, Appli
21 70 100.0 585 14 US-10-414-469-2 Sequence 2, Appli
22 70 100.0 585 14 US-10-413-831-1 Sequence 1, Appli
23 70 100.0 585 14 US-10-413-831-2 Sequence 2, Appli
24 70 100.0 585 15 US-10-413-832-1 Sequence 1, Appli
25 70 100.0 585 15 US-10-413-832-2 Sequence 2, Appli
26 70 100.0 585 15 US-10-414-386-1 Sequence 1, Appli
27 70 100.0 585 15 US-10-414-386-2 Sequence 2, Appli
28 70 100.0 585 15 US-10-433-675A-11 Sequence 2, Appli
29 70 100.0 585 15 US-10-432-262-26 Sequence 11, Appl
30 70 100.0 604 10 US-09-984-010-7 Sequence 26, Appl
31 70 100.0 609 10 US-09-919-039-370 Sequence 370, App
32 70 100.0 609 12 US-10-609-346-12 Sequence 12, Appl
33 70 100.0 609 13 US-10-153-064-7 Sequence 7, Appli
34 70 100.0 609 14 US-10-153-604A-7 Sequence 7, Appli
35 70 100.0 609 14 US-10-365-623-23 Sequence 23, Appl
36 70 100.0 610 9 US-09-984-186-2 Sequence 2, Appli
37 70 100.0 610 14 US-10-237-667-2 Sequence 2, Appli
38 70 100.0 610 14 US-10-237-708-2 Sequence 2, Appli
39 70 100.0 610 14 US-10-237-866-2 Sequence 2, Appli
40 70 100.0 610 14 US-10-237-871-2 Sequence 2, Appli
41 70 100.0 610 14 US-10-237-624-2 Sequence 2, Appli
42 70 100.0 616 12 US-10-433-108-13 Sequence 13, Appl
43 70 100.0 624 12 US-10-433-108-16 Sequence 16, Appl
44 70 100.0 631 12 US-10-433-108-14 Sequence 14, Appl
45 70 100.0 640 12 US-10-433-108-15 Sequence 15, Appl

```

#### ALIGNMENTS

```

RESULT 1
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-24

```

```

Query Match 100.0%; Score 70; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

```

```

RESULT 2
US-10-074-956-27
; Sequence 27, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956

```

```

XX 16-JUL-2001; 2001WO-US022263.
XX
XX 14-JUL-2000; 2000US-0218381P.
XX 18-AUG-2000; 2000US-0226382P.
XX 06-OCT-2000; 2000US-0238380P.
XX 29-DEC-2000; 2000US-0258764P.
XX 14-JUN-2001; 2001US-0298317P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
XX
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
XX stimulating hormone concatamer or its analog, for treating inflammatory
XX or autoimmune disorders.
XX
XX Example 2; Page 4-5; 89pp; English.
XX
XX The present invention relates to a nucleic acid comprising a sequence
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating
XX hormone (MSH) concatamer. The sequences are useful for treating an
XX individual suffering from, or at risk of, a disorder of the immune system
XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
XX present sequence is a protein described in the exemplification of the
XX invention
XX
XX Sequence 241 AA;
XX
XX Query Match 100.0%; Score 70; DB 5; Length 241;
XX Best Local Similarity 100.0%; Pred. No. 0.0001;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TVATLRETYGEMAD 14
XX |||||
XX Db 100 TVATLRETYGEMAD 113
XX
XX Search completed: April 19, 2004, 11:51:17
XX Job time : 9.23453 secs

```

DE Novel human secreted protein #365.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200179449-A2.  
 XX  
 XX 25-OCT-2001.  
 XX  
 XX 16-APR-2001; 2001WO-US008656.  
 XX  
 XX 18-APR-2000; 2000US-00552929.  
 XX  
 XX 26-JAN-2001; 2001US-00770160.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX WPI; 2001-611725/70.  
 DR  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 PT  
 XX  
 PS Claim 20; Page 205; 765pp; English.  
 XX  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 214 AA;  
 Query Match 100.0%; Score 70; DB 4; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVATLRETYGEMAD 14  
 Db 112 TVATLRETYGEMAD 125  
 RESULT 14  
 AA017051  
 ID AA017051 standard; protein; 236 AA.  
 XX  
 AC AA017051;  
 XX  
 XX 29-MAY-2002 (first entry)  
 DT  
 XX Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.  
 DE  
 XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;  
 KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
 XX

KW diabetes; uveitis; coeliac disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200206316-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 XX 16-JUL-2001; 2001WO-US022263.  
 XX  
 XX 14-JUL-2000; 2000US-0218381P.  
 XX  
 PR 18-AUG-2000; 2000US-0226382P.  
 PR  
 PR 06-OCT-2000; 2000US-0238380P.  
 PR  
 PR 29-DEC-2000; 2000US-0258764P.  
 PR  
 PR 14-JUN-2001; 2001US-0298317P.  
 XX  
 PA (ZYCO-) ZYCO INC.  
 XX  
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 PI  
 XX WPI; 2002-195801/25.  
 DR  
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 PT stimulating hormone concatamer or its analog, for treating inflammatory  
 PT or autoimmune disorders.  
 PT  
 XX  
 PS Example 2; Page 48; 89pp; English.  
 XX  
 CC The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 CC hormone (MSH) concatamer. The sequences are useful for treating an  
 CC individual suffering from, or at risk of, a disorder of the immune system  
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 CC present sequence is a peptide described in the exemplification of the  
 CC invention  
 XX  
 SQ Sequence 236 AA;  
 Query Match 100.0%; Score 70; DB 5; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVATLRETYGEMAD 14  
 Db 100 TVATLRETYGEMAD 113  
 RESULT 15  
 AA016984  
 ID AA016984 standard; protein; 241 AA.  
 XX  
 AC AA016984;  
 XX  
 XX 29-MAY-2002 (first entry)  
 DT  
 XX Alpha-MSH construct protein fragment SEQ ID NO: 59.  
 DE  
 XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;  
 KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
 KW diabetes; uveitis; coeliac disease.  
 XX  
 XX Unidentified.  
 XX  
 PN W0200206316-A2.  
 XX  
 XX 24-JAN-2002.



Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
Db 76 TVATLRETYGEMAD 89

RESULT 11  
ABU10022  
ID ABU10022 standard; protein; 195 AA.  
XX  
AC ABU10022;  
XX  
DT 31-JUL-2003 (first entry)  
XX  
DE Human serum albumin residues 1-195.  
XX  
KW Bladder disorder; cytostatic; antiinflammatory; immune response;  
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
KW bladder cancer; tumour; interstitial cystitis; inflammation;  
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;  
KW serum albumin.  
XX  
OS Homo sapiens.  
XX  
XX US2002193332-A1.  
XX  
PD 19-DEC-2002.  
XX  
PF 12-FEB-2002; 2002US-00074956.  
XX  
PR 12-FEB-2001; 2001US-0268175P.  
XX  
XX (HEDL/) HEDLEY M L.  
XX  
PI Hedley ML;  
XX  
XX WPI; 2003-447327/42.  
XX  
PT Modulating immune responses in a mammal with a bladder disorder e.g.  
PT bladder cancer, by administering nucleic acids comprising un-methylated  
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
PT the mammal.  
XX  
PS Example 2; Page 9; 17pp; English.  
XX

CC The invention describes a method of modulating an immune response in a  
CC mammal, comprising identifying a mammal that has or is at risk for having  
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated  
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;  
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
CC method is useful for modulating immune response in a mammal having a  
CC bladder disorder, where administration of (N1) results in an amelioration  
CC of one or more symptoms of the disorder. Preferably, the method is useful  
CC for modulating immune response in a mammal having bladder cancer (where  
CC administration of (N1) results in a decrease in tumour size or activity),  
CC or for modulating immune response in a mammal having interstitial  
CC cystitis (where administration of (N1) results in a modulation of the  
CC immune response from Th2 response to a Th1 response). The method is also  
CC useful for modulating immune response in a mammal having bladder disorder  
CC that is characterised by inflammation which is associated with symptoms  
CC of interstitial cystitis or associated with a disruption of the integrity  
CC of the bladder lining. This is the amino acid sequence of human serum  
CC albumin residues 1-195 that can be used in the creation of melanocyte  
CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the  
CC fusion protein when expressed in mammalian cells  
XX  
SQ Sequence 195 AA;

Query Match 100.0%; Score 70; DB 7; Length 195;  
Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
Db 76 TVATLRETYGEMAD 89

RESULT 12  
AAU29874  
ID AAU29874 standard; protein; 204 AA.  
XX  
AC AAU29874;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Yeast codon-biased recombinant HSA protein fragment HSA-I.  
XX  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX CN1239103-A.  
XX  
XX -22-DEC-1999.  
XX  
XX 17-JUN-1998; 98CN-00102506.  
XX  
XX 17-JUN-1998; 98CN-00102506.  
XX  
XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
XX  
XX Li S, Lu D;  
XX  
XX WPI; 2000-351198/31.  
XX  
XX N-PSDB; AAA10092.  
XX  
XX Process for preparing recombinant human serum albumin comprising yeast  
XX biased sex codons - uses a recombinant DNA technique.  
XX  
XX Example 1; Fig 3; 44pp; Chinese.  
XX

CC The method relates to a method of recombinantly producing human serum  
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as  
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA  
CC technology. Each HSA fragment was synthesised from overlapping  
CC oligonucleotide fragments that were extended. This sequence represents  
CC the sequence of the HSA fragment HSA-I encoded by the human gene with a  
CC yeast codon bias. The invention also covers a recombinant expression  
CC vector, yeast host cells carrying the recombinant expression vector and  
CC the process for producing human serum albumin in the yeast host cell,  
CC especially in secretory mode  
XX  
XX Sequence 204 AA;

Query Match 100.0%; Score 70; DB 3; Length 204;  
Best Local Similarity 100.0%; Pred. No. 8.3e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
Db 83 TVATLRETYGEMAD 96

RESULT 13  
AAU29874  
ID AAU29874 standard; protein; 214 AA.  
XX  
AC AAU29874;  
XX  
DT 18-DEC-2001 (first entry)  
XX

CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 156 AA;

Query Match 100.0%; Score 70; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | |  
Db 31 TVATLRETYGEMAD 44

RESULT 9  
AAU33271  
ID AAU33271 standard; protein; 156 AA.  
AC AAU33271;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
XX Novel human secreted protein #3762.  
XX  
XX Human; vaccination; gene therapy; nutritional supplement;  
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX WO200179449-A2.  
XX  
XX 25-OCT-2001.  
XX  
XX 16-APR-2001; 2001WO-US008656.  
XX  
XX 18-APR-2000; 2000US-00552929.  
XX  
XX 26-JAN-2001; 2001US-00770160.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-611725/70.  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
XX Claim 20; Page 751; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 156 AA;

Query Match 100.0%; Score 70; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | |  
Db 31 TVATLRETYGEMAD 44

RESULT 10  
AAO17048  
ID AAO17048 standard; protein; 195 AA.  
AC AAO17048;  
XX  
XX 29-MAY-2002 (first entry)  
XX  
XX Human serum albumin (1-195) SEQ ID NO: 56.

XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;  
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;  
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
KW diabetes; uveitis; coeliac disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200206316-A2.  
XX  
XX 24-JAN-2002.  
XX  
XX 16-JUL-2001; 2001WO-US022263.  
XX  
XX 14-JUL-2000; 2000US-0218381P.  
XX  
XX 18-AUG-2000; 2000US-0226382P.  
XX  
XX 06-OCT-2000; 2000US-0238380P.  
XX  
XX 29-DEC-2000; 2000US-0258764P.  
XX  
XX 14-JUN-2001; 2001US-0298317P.  
XX  
XX (ZYCO-) ZYCO INC.

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
XX WPI; 2002-195801/25.  
XX  
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
PT stimulating hormone concatamer or its analog, for treating inflammatory  
PT or autoimmune disorders.  
XX  
XX Example 2; Page 46; 89pp; English.

XX The present invention relates to a nucleic acid comprising a sequence  
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating  
XX hormone (MSH) concatamer. The sequences are useful for treating an  
XX individual suffering from, or at risk of, a disorder of the immune system  
XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
XX present sequence is a protein described in the exemplification of the  
XX invention  
XX  
XX Sequence 195 AA;

Query Match 100.0%; Score 70; DB 5; Length 195;  
Best Local Similarity 100.0%; Pred. No. 7.9e-05;

XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX N-PSDB; AA182561.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX Claim 20; SEQ ID NO 16522; 1399pp + Sequence Listing; English.  
 XX The invention relates to human polynucleotides (AA179941-AA193841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 124 AA;  
 Query Match 100.0%; Score 70; DB 4; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVATLRETYGENAD 14  
 DB 96 TVATLRETYGENAD 109  
 RESULT 7  
 AAU29708  
 ID AAU29708 standard; protein; 134 AA.  
 XX AAU29708;  
 XX 18-DEC-2001 (first entry)  
 XX Novel human secreted protein #199.  
 XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX Homo sapiens.  
 XX WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008656.  
 XX 18-APR-2000; 2000US-00552929.  
 XX 26-JAN-2001; 2001US-00770160.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX Claim 20; Page 183; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX SQ Sequence 134 AA;  
 Query Match 100.0%; Score 70; DB 4; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TVATLRETYGENAD 14  
 DB 7 TVATLRETYGENAD 20  
 RESULT 8  
 AAU33073  
 ID AAU33073 standard; protein; 156 AA.  
 XX AAU33073;  
 XX 18-DEC-2001 (first entry)  
 XX Novel human secreted protein #3564.  
 XX Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX Homo sapiens.  
 XX WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US008656.  
 XX 18-APR-2000; 2000US-00552929.  
 XX 26-JAN-2001; 2001US-00770160.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX Claim 20; Page 704; 765pp; English.  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or

```

XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 18316.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI84355.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 18316; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 123 AA;
XX Query Match 100.0%; Score 70; DB 4; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 98 TVATLRETYGEMAD 111

RESULT 5
AAO04423
ID AAO04423 standard; protein; 123 AA.
XX AC AAO04423;
XX OS Homo sapiens.
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 18315.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX

```

```

OS Homo sapiens.
XX WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI84354.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 18315; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 123 AA;
XX Query Match 100.0%; Score 70; DB 4; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 99 TVATLRETYGEMAD 112

RESULT 6
AAO02630
ID AAO02630 standard; protein; 124 AA.
XX AC AAO02630;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 16522.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.

```

CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 114 AA;

Query Match 100.0%; Score 70; DB 4; Length 114;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 101 TVATLRETYGEMAD 114

RESULT 2  
AAO04446  
ID AAO04446 standard; protein; 118 AA.

XX AC AAO04446;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 18338.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI84377.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 18338; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation. Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 118 AA;

Query Match 100.0%; Score 70; DB 4; Length 118;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 98 TVATLRETYGEMAD 111

RESULT 3

AAO01996

ID AAO01996 standard; protein; 119 AA.

XX AC AAO01996;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 15888.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI81927.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 15888; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation. Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 119 AA;

Query Match 100.0%; Score 70; DB 4; Length 119;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
Db 98 TVATLRETYGEMAD 111

RESULT 4

AAO04424

ID AAO04424 standard; protein; 123 AA.

XX AC AAO04424;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 8.23453 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_76\_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_29Jan04:.\*  
1: Geneseqpl1980s:.\*  
2: Geneseqpl1990s:.\*  
3: Geneseqpl2000s:.\*  
4: Geneseqpl2001s:.\*  
5: Geneseqpl2002s:.\*  
6: Geneseqpl2003as:.\*  
7: Geneseqpl2003bs:.\*  
8: Geneseqpl2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID    | Description        |
|------------|-------|---------|--------------|----------|--------------------|
| 1          | 70    | 100.0   | 114          | AAO09588 | AAO09588 Human pol |
| 2          | 70    | 100.0   | 118          | AAO04446 | AAO04446 Human pol |
| 3          | 70    | 100.0   | 119          | AAO01996 | AAO01996 Human pol |
| 4          | 70    | 100.0   | 123          | AAO04424 | AAO04424 Human pol |
| 5          | 70    | 100.0   | 123          | AAO04423 | AAO04423 Human pol |
| 6          | 70    | 100.0   | 124          | AAO02630 | AAO02630 Human pol |
| 7          | 70    | 100.0   | 134          | AAU29708 | AAU29708 Novel hum |
| 8          | 70    | 100.0   | 156          | AAU33073 | AAU33073 Novel hum |
| 9          | 70    | 100.0   | 156          | AAU33271 | AAU33271 Novel hum |
| 10         | 70    | 100.0   | 195          | AAO17048 | AAO17048 Human ser |
| 11         | 70    | 100.0   | 195          | ABU10022 | ABU10022 Human ser |
| 12         | 70    | 100.0   | 204          | AAU83947 | AAU83947 Yeast cod |
| 13         | 70    | 100.0   | 214          | AAU29874 | AAU29874 Novel hum |
| 14         | 70    | 100.0   | 236          | AAO17051 | AAO17051 Human alb |
| 15         | 70    | 100.0   | 241          | AAO16984 | AAO16984 Alpha-MSH |
| 16         | 70    | 100.0   | 241          | ABU10025 | ABU10025 Alpha-MSH |
| 17         | 70    | 100.0   | 242          | AAO16985 | AAO16985 Alpha-MSH |
| 18         | 70    | 100.0   | 244          | AAO16986 | AAO16986 Alpha-MSH |
| 19         | 70    | 100.0   | 245          | AAO16987 | AAO16987 Alpha-MSH |
| 20         | 70    | 100.0   | 245          | AAO16988 | AAO16988 Alpha-MSH |
| 21         | 70    | 100.0   | 268          | AAO16989 | AAO16989 Alpha-MSH |
| 22         | 70    | 100.0   | 268          | ABU10026 | ABU10026 Alpha-MSH |
| 23         | 70    | 100.0   | 303          | AAU14178 | AAU14178 Human ser |
| 24         | 70    | 100.0   | 327          | AAU32564 | AAU32564 Novel hum |
| 25         | 70    | 100.0   | 327          | AAU29942 | AAU29942 Novel hum |

|    |    |       |     |   |          |                    |
|----|----|-------|-----|---|----------|--------------------|
| 26 | 70 | 100.0 | 327 | 4 | AAU32995 | AAU32995 Novel hum |
| 27 | 70 | 100.0 | 327 | 4 | AAU33287 | AAU33287 Novel hum |
| 28 | 70 | 100.0 | 373 | 1 | AAU90387 | AAU90387 N-termina |
| 29 | 70 | 100.0 | 388 | 1 | AAU90389 | AAU90389 N-termina |
| 30 | 70 | 100.0 | 389 | 1 | AAU90390 | AAU90390 N-termina |
| 31 | 70 | 100.0 | 390 | 1 | AAU90391 | AAU90391 N-termina |
| 32 | 70 | 100.0 | 401 | 4 | AAU29876 | AAU29876 Novel hum |
| 33 | 70 | 100.0 | 407 | 1 | AAU90392 | AAU90392 N-termina |
| 34 | 70 | 100.0 | 500 | 7 | ADD32019 | ADD32019 Heterolog |
| 35 | 70 | 100.0 | 584 | 6 | ABG72381 | ABG72381 Mature hu |
| 36 | 70 | 100.0 | 585 | 1 | AAU93344 | AAU93344 Sequence  |
| 37 | 70 | 100.0 | 585 | 1 | AAU90388 | AAU90388 Mature hu |
| 38 | 70 | 100.0 | 585 | 1 | AAU91422 | AAU91422 Human nor |
| 39 | 70 | 100.0 | 585 | 2 | AAU05318 | AAU05318 Human ser |
| 40 | 70 | 100.0 | 585 | 2 | AAU08457 | AAU08457 Human ser |
| 41 | 70 | 100.0 | 585 | 2 | AAU26207 | AAU26207 Human ser |
| 42 | 70 | 100.0 | 585 | 2 | AAU26362 | AAU26362 Synthetic |
| 43 | 70 | 100.0 | 585 | 2 | AAU20029 | AAU20029 Human ser |
| 44 | 70 | 100.0 | 585 | 2 | AAU80301 | AAU80301 Human ser |
| 45 | 70 | 100.0 | 585 | 2 | AAO20111 | AAO20111 HSA prote |

## ALIGNMENTS

### RESULT 1

AAO09588  
ID AAO09588 standard; protein; 114 AA.  
XX  
AC AAO09588;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 23480.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US004927.  
XX  
PR 28-FEB-2000; 2000US-00515126.  
XX  
PA 18-MAY-2000; 2000US-00577409.  
XX  
(HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
XX  
N-PSDB; AA189519.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
PS Claim 20; SEQ ID NO 23480; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

RL Nat. Biotechnol. 21:526-531 (2003).  
 DR EMBL; AP005037; BAC71712.1; -.  
 DR InterPro; IPR000437; Prok\_lipoprot S.  
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 169 AA; 17778 MW; 9D881D18DB331D19 CRC64;

Query Match 57.1%; Score 40; DB 16; Length 169;  
 Best Local Similarity 61.5%; Pred. No. 43;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRETYGENAD 14  
 |||||  
 Db 41 VMFLREGYGLAD 53

Search completed: April 19, 2004, 12:00:01  
 Job time : 7.40351 secs

```

RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640432; CAE38219.1; -.
KW Complete proteome.
SQ SEQUENCE 281 AA; 30591 MW; C97F4A51E40E3315 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 281;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
    |||||
Db 266 ATLRLYGELED 277

RESULT 13
CO1889 PRELIMINARY; PRT; 354 AA.
AC O01889;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE R08F11.4 protein.
GN R08F11.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx P.;
RT "The sequence of C. elegans cosmid R08F11."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003385; AAB54246.1; -.
DR PIR; H89009; H89009
DR WormPep; R08F11.4; CE12586.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR InterPro; IPR000051; SAM_bind.
SQ SEQUENCE 354 AA; 39365 MW; CB422510DFA3E0A CRC64;

Query Match 58.6%; Score 41; DB 5; Length 354;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VATLRETYGEMA 13
    |||||
Db 280 VFTDRTYKGMA 291

RESULT 14
Q9A8L1 PRELIMINARY; PRT; 692 AA.
ID Q9A8L1
AC Q9A8L1;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Glycyl-tRNA synthetase, beta subunit.
GN CC1342.

```

```

OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Stepan M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uytterback I., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; A3005809; AAK23323.1; -.
DR PIR; G87415; G87415.
DR TIGR; CC1342; -.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0004820; F:glycine-tRNA ligase activity; IEA.
GO; GO:0006426; P:glycyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002311; trna_synth_2f.
DR InterPro; IPR006194; trna_synth_gly.
DR Pfam; PF02092; trna_synth_2f; 1.
DR PRINTS; PR01045; TRNASYNTHGB.
DR TIGRFAMs; TIGR00211; glyS; 1.
DR PROSITE; PFS0861; AA_TRNA_LIGASE_II_GLYAB; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 692 AA; 76060 MW; 3CDE4E730D097846 CRC64;

Query Match 58.6%; Score 41; DB 16; Length 692;
Best Local Similarity 50.0%; Pred. No. 13e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
    |||||
Db 673 TLATVRDAMQVAD 686

RESULT 15
Q82G98 PRELIMINARY; PRT; 169 AA.
ID Q82G98
AC Q82G98;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN SAV4000.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."

```



```

DB 01-MAR-2002 (TReMBLrel. 20, Created)
DB 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DB 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DB Hypothetical protein All5024.
DB ALU5024.
DB Anabaena sp. (strain PCC 7120).
DB Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
DB NCBI_TaxID=103690;
DB [1]
DB SEQUENCE FROM N.A.
DB MEDLINE=21595285; PubMed=11759840;
DB Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
DB Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
DB Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
DB Nakazaki N., Shampo S., Sugimoto M., Takazawa M., Yanada M.,
DB Yasuda M., Tabata S.;
DB "Complete genomic sequence of the filamentous nitrogen-fixing
DB cyanobacterium Anabaena sp. strain PCC 7120.";
DB DNA Res. 8:205-213(2001);
DB EMBL; AP003598; BAB76723.1; -.
DB PIR; AH2433; AH2433.
DB Hypothetical protein; Complete proteome.
DB SEQUENCE 252 AA; 29440 MW; 3ABDCE406345B50F CRC64;

Query Match 60.0%; Score 42; DB 16; Length 252;
Best Local Similarity 64.3%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TVATLRITYGEMAD 14
:|||||
Db 176 TIAIIRSTAHMMD 189

RESULT 10
QWIF5 PRELIMINARY; PRT; 271 AA.
AC QWIF5;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN B2896.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR Nat. Genet. 35:32-40(2003);
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGLID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC QW613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR Nat. Genet. 35:32-40(2003);
DR EMBL; EX640445; CAE33388.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29643 MW; F5E8C19A837AB0E CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||

```

```

Db 256 ATLRLYGLID 267

RESULT 11
QWYTM6 PRELIMINARY; PRT; 271 AA.
AC QWYTM6;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative exported protein.
GN BPI292.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR Nat. Genet. 35:32-40(2003);
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGLID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC QW613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR Nat. Genet. 35:32-40(2003);
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGLID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC QW613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR Nat. Genet. 35:32-40(2003);
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||

```

Best Local Similarity 61.5%; Pred. No. 13; Mismatches 2; Indels 0; Gaps 0;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 VATLRETYGEMAD 14  
DB 69 IPNLRENYGELAD 81

RESULT 6  
Q8C7H3 PRELIMINARY; PRT; 608 AA.  
AC Q8C7H3;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Albumin 1.  
DE ALB1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=liver;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The PANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK041748; BAC31052.1; -.  
DR MGI; MGI:87991; Alb1.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR00264; Serum\_albumin.  
DR Pfam; PF00273; transport prot.; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SMC0103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
DR SEQUENCE 608 AA; 68722 MW; 292P600EED3A61B4 CRC64;

Query Match 65.7%; Score 46; DB 11; Length 608;  
Best Local Similarity 61.5%; Pred. No. 14;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 VATLRETYGEMAD 14  
DB 101 IPNLRENYGELAD 113

RESULT 7  
Q8C9N0 PRELIMINARY; PRT; 116 AA.  
AC Q8C9N0;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Pan hematopoietic expression.  
DE PHEMX.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Thymus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The PANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).  
DR EMBL; AK041748; BAC31052.1; -.  
DR MGI; MGI:1350360; PHEMX.  
DR GO; GO:0005622; C:intracellular; IEA.  
SQ SEQUENCE 116 AA; 12587 MW; 7638C7C4CFD2CDDC CRC64;

Query Match 60.0%; Score 42; DB 11; Length 116;  
Best Local Similarity 61.5%; Pred. No. 13; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TVATLRETYGEMA 13  
DB 80 TIATVREAHGLMA 92

RESULT 8  
Q7TNQ7 PRELIMINARY; PRT; 201 AA.  
AC Q7TNQ7;  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Salivary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Krzyzinski A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzyzinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Salivary gland;  
RA Strausberg R.;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; SC055858; AAH55858.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 201 AA; 22722 MW; 09421AC54204F79B CRC64;

Query Match 60.0%; Score 42; DB 11; Length 201;  
Best Local Similarity 61.5%; Pred. No. 23;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TVATLRETYGEMA 13  
DB 53 TIATVREAHGLMA 65

RESULT 9  
Q8YMB0 PRELIMINARY; PRT; 252 AA.  
ID Q8YMB0  
AC Q8YMB0;

```

ID QYVSG3 PRELIMINARY; PRT; 584 AA.
AC QYVSG3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ487677; CAD32275.1; -.
DR NON_TER 1
FT SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
SQ SEQUENCE 584 AA; 65908 MW; 81.4%; Score 57; DB 6; Length 584;
Query Match 81.4%; Score 57; DB 6; Length 584;
Best Local Similarity 78.6%; Pred. No. 0.15;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TVATLRITYGEMAD 14
Db 76 TVASLRDXYGEMAD 89

RESULT 3
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E34 CRC64;
Query Match 70.0%; Score 49; DB 5; Length 608;
Best Local Similarity 61.5%; Pred. No. 4.2;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VATLRITYGEMAD 14
Db 101 IPTLRDSYGLAD 113

RESULT 4

```

```

Q8CG74 PRELIMINARY; PRT; 205 AA.
ID Q8CG74;
AC Q8CG74;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvetOla;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277794; CAC81903.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 1.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
DR NON_TER 205
FT SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;
SQ SEQUENCE 205 AA; 23624 MW; 65.7%; Score 46; DB 11; Length 205;
Query Match 65.7%; Score 46; DB 11; Length 205;
Best Local Similarity 61.5%; Pred. No. 4.5;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VATLRITYGEMAD 14
Db 101 IPNLRENYGELAD 113

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7;
AC Q8C7C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR NON_TER 1
FT SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;
SQ SEQUENCE 576 AA; 65002 MW; 65.7%; Score 46; DB 11; Length 576;
Query Match 65.7%; Score 46; DB 11; Length 576;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 5.40351 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_76\_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 70    | 100.0       | 396    | 4 Q8IU7   | Q8IU7 homo sapien  |
| 2          | 57    | 81.4        | 584    | 6 Q7YSG3  | Q7YSG3 felis silve |
| 3          | 49    | 70.0        | 608    | 5 Q95VB7  | Q95VB7 schistosoma |
| 4          | 46    | 65.7        | 205    | 11 Q8CG74 | Q8CG74 mus musculu |
| 5          | 46    | 65.7        | 576    | 11 Q8C7C7 | Q8C7C7 mus musculu |
| 6          | 46    | 65.7        | 608    | 11 Q8C7H3 | Q8C7H3 mus musculu |
| 7          | 42    | 60.0        | 116    | 11 Q8C9N0 | Q8C9N0 mus musculu |
| 8          | 42    | 60.0        | 201    | 11 Q7TNQ7 | Q7TNQ7 mus musculu |
| 9          | 42    | 60.0        | 252    | 16 Q8YMB0 | Q8YMB0 anabaena sp |
| 10         | 42    | 60.0        | 271    | 16 Q7WIF5 | Q7WIF5 bordetella  |
| 11         | 42    | 60.0        | 271    | 16 Q7VYM6 | Q7VYM6 bordetella  |
| 12         | 42    | 60.0        | 281    | 16 Q7W6I3 | Q7W6I3 bordetella  |
| 13         | 41    | 58.6        | 354    | 5 O01889  | O01889 caenorhabdi |
| 14         | 41    | 58.6        | 692    | 16 Q9A8L1 | Q9A8L1 caulobacter |
| 15         | 40    | 57.1        | 169    | 16 Q82G98 | Q82G98 streptomyce |
| 16         | 40    | 57.1        | 270    | 16 Q8F5F7 | Q8F5F7 leptospira  |

```

17 40 57.1 314 16 Q83BP5
18 40 57.1 330 16 Q97R21
19 40 57.1 330 16 Q8E6D5
20 40 57.1 330 16 Q8E6D5
21 40 57.1 330 16 Q8E6D5
22 40 57.1 563 5 Q964I3
23 40 57.1 585 16 Q9RR13
24 40 57.1 924 2 Q9ZIG3
25 40 57.1 924 2 Q9X582
26 39.5 56.4 202 2 Q9ZGF5
27 39 55.7 157 5 Q90743
28 39 55.7 176 16 Q8DL82
29 39 55.7 179 16 Q9CELO
30 39 55.7 250 5 Q8SRK2
31 39 55.7 443 2 Q9L3B5
32 39 55.7 443 2 Q53162
33 39 55.7 444 2 Q9X7L9
34 39 55.7 445 16 Q8P181
35 39 55.7 541 16 Q9RKD1
36 39 55.7 633 12 Q92095
37 39 55.7 1049 2 Q9XBP6
38 39 55.7 3898 12 Q92365
39 39 55.7 3898 12 Q92872
40 38 54.3 91 15 Q8XSP0
41 38 54.3 124 17 Q8ZVT1
42 38 54.3 251 16 Q8XFB9
43 38 54.3 308 16 Q8XRG7
44 38 54.3 318 2 Q7WZ93
45 38 54.3 378 16 Q8FPK8

```

#### ALIGNMENTS

#### RESULT 1

```

Q8IU7
ID Q8IU7 PRELIMINARY; PRT; 396 AA.
AC Q8IU7;
DT 01-VAR-2003 (TRENBLrel. 23, Created)
DT 01-NAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005635; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

```

Query Match. 100.0%; Score 70; DB 4; Length 396;

Best Local Similarity 100.0%; Pred. No. 0.00047;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14

Db 100 TVATLRETYGEMAD 113

#### RESULT 2

Q7YSG3

Query Match 57.1%; Score 40; DB 1; Length 409;  
 Best Local Similarity 70.0%; Pred. No. 7.8;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAIURETYGE 11  
 : |||: |||  
 Db 217 INTLRKTYGE 226

Search completed: April 19, 2004, 11:52:48  
 Job time : 3.02124 secs

DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Glycine dehydrogenase [decarboxylating], mitochondrial precursor  
 DE (EC 1.4.4.2) (Glycine decarboxylase) (Glycine cleavage system P-  
 protein)  
 DE GCV2 OR GSD2 OR YMR189W OR YN5646.01.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_taxid=932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96120340; PubMed=7498764;  
 RA Sinclair D.A., Dawes I.W.;  
 RT "Genetics of the synthesis of serine from glycine and the utilization  
 of glycine as sole nitrogen source by *Saccharomyces cerevisiae*.";  
 RL Genetics 140:1213-1222(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=S288c / AB972;  
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
 RA Jagals K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,  
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;  
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome  
 XIII.";  
 RL Nature 387:90-93(1997).  
 CC -!- FUNCTION: The glycine cleavage system catalyzes the degradation of  
 glycine. The P-protein binds the alpha-amino group of glycine  
 through its pyridoxal phosphate cofactor; CO(2) is released and  
 the remaining methylamine moiety is then transferred to the  
 liponamide cofactor of the H-protein.  
 CC -!- CATALYTIC ACTIVITY: Glycine + lipoylprotein = S-  
 aminomethylidihydrolipoylprotein + CO(2).  
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
 CC -!- SIMILARITY: Belongs to the gcvP family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U20641; AAB18933.1; -.  
 DR EMBL; Z47815; CA887810.1; -.  
 DR PIR; S50917; S50917.  
 DR GenOnline; 142863; -.  
 DR SGD; S0004801; GCV2.  
 DR GO; GO:0004375; F:glycine dehydrogenase (decarboxylating) act. . . ; IEP.  
 DR GO; GO:0006730; P:one-carbon compound metabolism; IGT.  
 DR InterPro; IPR003437; GDC-P.  
 DR Pfam; PF02347; GDC-P; 1.  
 DR TIGRFAMs; TIGR00461; gcvP; 1.  
 DR Oxidoreductase; Mitochondrion; Pyridoxal phosphate; Transit peptide.  
 KW TRANSIT 1 ? MITOCHONDRION (POTENTIAL).  
 FT CHAIN 773 1034  
 FT BINDING 773 773 GLYCINE DEHYDROGENASE [DECARBOXYLATING].  
 FT BINDING 1034 1034 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 SQ SEQUENCE 1034 AA; 114451 MW; F4D52642B0BDA041 CRC64;  
 Query Match 58.6%; Score 41; DB 1; Length 1034;  
 Best Local Similarity 66.7%; Pred. No. 14;  
 Matches 8; Conservative 2; Mismatches 0; Gaps 0;  
 QY 1 TVATLREYIGEM 12  
 DB 1006 TVARLDYIGDM 1017  
 RESULT 15

ODO2\_HABIN STANDARD; PRT; 409 AA.  
 ID P45302;  
 AC DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Dihydrolipoamide succinyltransferase component of 2-oxoglutarate  
 DE dehydrogenase complex (EC 2.3.1.61) (E2).  
 GN SUCB OR H1661.  
 OS Haemophilus influenzae.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 CC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kierlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*  
 Rd.";  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE  
 OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT  
 CONTAINS MULTIPLE COPIES OF 3 ENZYMIC COMPONENTS: 2-OXOGLUTARATE  
 DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND  
 LIPOAMIDE DEHYDROGENASE (E3) (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: Succinyl-CoA + dihydrolipoamide = CoA + S-  
 succinylidihydrolipoamide.  
 CC -!- COFACTOR: Contains 1 covalently bound lipoyl cofactor (By  
 similarity).  
 CC -!- PATHWAY: Tricarboxylic acid cycle.  
 CC -!- SUBUNIT: Forms a 24-polypeptide structural core with octahedral  
 symmetry (By similarity).  
 CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.  
 CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U32839; AAC23307.1; -.  
 DR PIR; D64135; D64135.  
 DR HSSP; P07016; 1C4T.  
 DR TIGR; H11661; -.  
 DR InterPro; IPR001078; 2Oxoacid dh.  
 DR InterPro; IPR000089; Biotin lipoyl.  
 DR InterPro; IPR004367; E3\_binding.  
 DR InterPro; IPR003016; Lipoy\_BS.  
 DR InterPro; IPR006255; SUCB.  
 DR Pfam; PF00198; 2-oxoacid dh; 1.  
 DR Pfam; PF02817; e3 binding; 1.  
 DR Pfam; PF00364; biotin lipoyl; 1.  
 DR ProDom; PD001115; 2Oxoacid dh; 1.  
 DR TIGRFAMs; TIGR01347; sucB; 1.  
 DR PROSITE; PS00189; LIPOYL; 1.  
 DR Tricarboxylic acid cycle; transferase; Acyltransferase; Lipoyl;  
 KW Complete proteome.  
 FT BINDING 43 43 LIPOYL (POTENTIAL).  
 FT ACT\_SITE 380 380 POTENTIAL.  
 FT ACT\_SITE 384 384 POTENTIAL.  
 SQ SEQUENCE 409 AA; 45162 MW; 3FBF62BC17433839 CRC64;

```

FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT VARIANT 262 262
FT CONFLICT 174 174
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

Query Match 60.0%; Score 42; DB 1; Length 608;
Best Local Similarity 53.8%; Pred. No. 5.1;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VATIRETYGEMAD 14
Db 101 IPLKRDNYGELAD 113

RESULT 13
ALBU CHICK STANDARD; PRT; 615 AA.
AC P1912;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
GN ALB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Casady A.I., Salkild C.K., Baverstock P., Wallace J.C.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=83161037; PubMed=6187737;
RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
RT "The 5' noncoding and flanking regions of the avian very low density
RT apolipoprotein II and serum albumin genes. Homologies with the egg
RT white protein genes.";
RL J. Biol. Chem. 258:4556-4564 (1983).
RN [3]
RN SEQUENCE OF 19-30.
RX MEDLINE=78019943; PubMed=911327;
RA Rosen A.M., Geller D.M.;
RT "Chicken microsomal albumin: amino terminal sequence of chicken
RT proalbumin.";
RL Blochem. Biophys. Res. Commun. 78:1060-1066 (1977).
RN [4]
RN ALLERGENIC PROPERTIES.
RX MEDLINE=21381307; PubMed=11488669;
RA Quirce S., Maranon F., Umpierrez A., de las Heras M.,
RA Fernandez-Caldas E., Sastre J.;
RT "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
RT and food allergen implicated in the bird-egg syndrome.";
RL Allergy 56:754-762 (2001).

```

-!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
 -!- SUBCELLULAR LOCATION: Secreted.  
 -!- TISSUE SPECIFICITY: Plasma.  
 -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE. Partially heat-labile allergen that may cause both respiratory and food-allergy symptoms in patients with the bird-egg syndrome.  
 -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X60688; CAA43098.1; -;  
 EMBL; V00381; CAA23680.1; -;  
 DR PIR; S15571; ABCHS.  
 DR HSP; P02768; 1E7B.  
 DR InterPro: IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT SIGNAL 1 18  
 FT PROPEP 19 23  
 FT CHAIN 24 615 SERUM ALBUMIN.  
 FT DOMAIN 24 209 ALBUMIN 1.  
 FT DOMAIN 216 401 ALBUMIN 2.  
 FT DOMAIN 408 599 ALBUMIN 3.  
 FT METAL 30 30 COPPER (BY SIMILARITY).  
 FT DISULFID 80 89 BY SIMILARITY.  
 FT DISULFID 102 118 BY SIMILARITY.  
 FT DISULFID 117 128 BY SIMILARITY.  
 FT DISULFID 152 197 BY SIMILARITY.  
 FT DISULFID 196 205 BY SIMILARITY.  
 FT DISULFID 228 274 BY SIMILARITY.  
 FT DISULFID 273 281 BY SIMILARITY.  
 FT DISULFID 293 307 BY SIMILARITY.  
 FT DISULFID 306 317 BY SIMILARITY.  
 FT DISULFID 344 389 BY SIMILARITY.  
 FT DISULFID 388 397 BY SIMILARITY.  
 FT DISULFID 420 466 BY SIMILARITY.  
 FT DISULFID 465 476 BY SIMILARITY.  
 FT DISULFID 489 505 BY SIMILARITY.  
 FT DISULFID 504 515 BY SIMILARITY.  
 FT DISULFID 542 587 BY SIMILARITY.  
 FT DISULFID 586 595 BY SIMILARITY.  
 FT CARBOHYD 500 500 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 24 24 F -> M (IN REF. 3).  
 SQ SEQUENCE 615 AA; 69918 MW; E59E4BBAEC066C6 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 615;  
 Best Local Similarity 61.5%; Pred. No. 7.9;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VATIRETYGEMAD 14  
 Db 104 VEKLRDSYGAMAD 116

RESULT 14  
 GCSP\_YEAST  
 ID GCSP\_YEAST STANDARD; PRT; 1034 AA.  
 AC P49095;  
 DT 01-FEB-1996 (Rel. 33, Created)

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X12422; CAA30970.1; -;  
EMBL; M36787; AAA30988.1; -;  
PIR; S01382; ABPGS.  
HSP; P02768; 187H.  
InterPro; IPR000264; Serum albumin.  
Pfam; PF00273; transport\_prot; 3.  
PRINTS; PR00802; SERUMALBUMIN.  
ProDom; PD002486; Serum albumin; 1.  
SMART; SM00103; ALBUMIN; 3.  
PROSITE; PS00212; ALBUMIN; 3.  
Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
NON\_TER 1 1  
SIGNAL <1 16 BY SIMILARITY.  
PROPEP 17 22 BY SIMILARITY.  
CHAIN 23 605 SERUM ALBUMIN.  
DOMAIN 23 202 ALBUMIN 1.  
DOMAIN 209 394 ALBUMIN 2.  
DOMAIN 401 592 ALBUMIN 3.  
METAL 31 31 COPPER (BY SIMILARITY).  
DISULFID 75 84 BY SIMILARITY.  
DISULFID 97 113 BY SIMILARITY.  
DISULFID 112 123 BY SIMILARITY.  
DISULFID 145 190 BY SIMILARITY.  
DISULFID 189 198 BY SIMILARITY.  
DISULFID 221 267 BY SIMILARITY.  
DISULFID 266 274 BY SIMILARITY.  
DISULFID 286 300 BY SIMILARITY.  
DISULFID 299 310 BY SIMILARITY.  
DISULFID 337 382 BY SIMILARITY.  
DISULFID 381 390 BY SIMILARITY.  
DISULFID 413 459 BY SIMILARITY.  
DISULFID 458 469 BY SIMILARITY.  
DISULFID 482 498 BY SIMILARITY.  
DISULFID 497 508 BY SIMILARITY.  
DISULFID 535 580 BY SIMILARITY.  
DISULFID 579 588 BY SIMILARITY.  
CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).  
SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;  
Query Match 60.0%; Score 42; DB 1; Length 605;  
Best Local Similarity 53.8%; Pred. No. 5;  
Matches 7; Conservativity 4; Mismatches 2; Indels 0; Gaps 0;  
QY 2 VATLRETYGENAD 14  
D5 99 IPSLREHYGLAD 111  
RESULT 12  
ALBU RAT STANDARD; PRT; 608 AA.  
AC P02770; P11382;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].  
GN ALB.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]\_TaxID=10116;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81223722; PubMed=7017712;  
RA Sargent T.D., Yang M., Bonner J.;  
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";

Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).  
[2] SEQUENCE OF 1-38, AND PROCESSING.  
RX MEDLINE=77249557; PubMed=893447;  
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-  
piece. Analysis of the direct translation product of albumin  
messenger RNA.";  
RT J. Biol. Chem. 252:6846-6855(1977).  
[3] SEQUENCE OF 25-222.  
RX MEDLINE=78109429; PubMed=564345;  
RA Isemura S., Ikenaka T.;  
RT "Amino acid sequences of fragments I and II obtained by cyanogen  
bromide cleavage of rat serum albumin.";  
RL J. Biochem. 83:35-48(1978).  
[4] SEQUENCE OF 223-288 AND 572-608.  
RX MEDLINE=76260153; PubMed=956149;  
RA Isemura S., Ikenaka T.;  
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
the amino acid sequences of four fragments.";  
RL J. Biochem. 79:1183-1196(1976).  
[5] SEQUENCE OF 166-174.  
RX TISSUE-Plasma;  
MEDLINE=87194805; PubMed=2437111;  
RA Carraway R.E., Mitra S.P., Cochran D.E.;  
RT "Structure of a biologically active neurotensin-related peptide  
obtained from pepsin-treated albumin(s).";  
RL J. Biol. Chem. 262:5968-5973(1987).  
[6] COPPER-BINDING.  
RX MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
Cancer Res. 38:3483-3486(1978).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and  
blood flow (Potential).  
CC -!- SUBCELLULAR LOCATION: Plasma.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; V01222; CAA24532.1; -;  
PIR; A93872; ABRTS.  
HSP; P02768; 1E7B.  
InterPro; IPR000264; Serum albumin.  
Pfam; PF00273; transport\_prot; 3.  
PRINTS; PR00802; SERUMALBUMIN.  
ProDom; PD002486; Serum albumin; 1.  
SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24 SERUM ALBUMIN.  
FT CHAIN 25 608 NEUROTENSIN-RELATED PEPTIDE.  
FT PEPTIDE 166 174 ALBUMIN 1.  
FT DOMAIN 25 205 ALBUMIN 1.  
FT DOMAIN 212 387 ALBUMIN 2.  
FT DOMAIN 404 595 ALBUMIN 3.





CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR EMBL; AJ011413; CAA09617.1; -
DR EMBL; AK010025; BAB26650.1; -
DR EMBL; BC049971; AAA49971.1; -
DR EMBL; M16111; AAA37190.1; -
DR EMBL; X13060; CAA31458.1; -
DR F01; A05139; A05139.
DR HSSP; P02768; 1E7B.
DR SWISS-2DPAGE; P07724; MOUSE.
DR MGD; MGI:97991; Alb1.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CONFLICT 27 27
FT CONFLICT 33 33
FT CONFLICT 41 41
SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;

Query Match 65.7%; Score 46; DB 1; Length 608;
Best Local Similarity 61.5%; Pred. No. 0.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
DB 101 IPNLRENYGELAD 113

RESULT 9
ALBU RABIT
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL SUBMITTED (DEC-1994) to the EMBL/GenBank/DBAJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  

CC or send an email to license@isb-sib.ch).

```



```

RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RL "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
CC Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PD00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;
Query Match 81.4%; Score 57; DB 1; Length 608;
Best Local Similarity 78.6%; Pred No. 0.0078;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 TVATLRETYGEMAD 14
Db 100 TVASLRDKYGMAD 113
RESULT 7
ALBU_CANFA
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O77705; Q9TSZ4;
DT 1-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

```

DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweiger C., Speir W.R., Pandjaitan B., Valent P.,
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive
RT allergen.";
RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ133489; CAB64867.1; -.
DR EMBL; Y17737; CAA76841.1; -.
DR EMBL; S72946; AAB30434.1; -.
DR HSSP; P02768; 1E7B.
DR HSC-2DPAGE; P49822; DOG.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PD00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.

```

RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
RA Wu H.T., Huang M.C.;  
RT "The complete cDNA sequence of bovine serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 1-32.  
RX MEDLINE=80024278; PubMed=488109;  
RA McGillivray R.T.A., Chung D.W., Davie E.W.;  
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin.";  
RL Eur. J. Biochem. 98:477-485(1979).  
RN [6]  
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.  
RA Brown J.R.;  
RT "Structure of bovine serum albumin.";  
RL Fed. Proc. 34:591-591(1975).  
RN [7]  
RP REVISIONS TO 190-195.  
RA Brown J.R.;  
RT Submitted (APR-1975) to the PIR data bank.  
RN [8]  
RP SEQUENCE OF 402-433.  
RX MEDLINE=8203364; PubMed=7283978;  
RA Reed R.G., Putnam F.W., Peters T. Jr.;  
RT "Sequence of residues 400-403 of bovine serum albumin.";  
RL Biochem. J. 191:867-868(1980).  
RN [9]  
RP SEQUENCE OF 19-28.  
RX MEDLINE=7134075; PubMed=843354;  
RA Patterson J.E., Geller D.M.;  
RT "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin.";  
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
RN [10]  
RP SEQUENCE AND REVISIONS TO 118-119 AND 180.  
RX MEDLINE=91083649; PubMed=2260975;  
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;  
RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by EMS and Frit-FAB LC/MS.";  
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).  
RN [11]  
RP SEQUENCE OF 25-41.  
RX MEDLINE=88267456; PubMed=3389500;  
RA Haen J.C., Lin F., Tam M.F.;  
RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";  
RL Anal. Biochem. 170:1-8(1988).  
RN [12]  
RP SEQUENCE OF 437-451.  
RA Vilbois F.;  
RT Submitted (AUG-1998) to Swiss-Prot.  
RN [13]  
RP DISULFIDE BONDS.  
RA Brown J.R.;  
RT "Structure of serum albumin: disulfide bridges.";  
RL Fed. Proc. 33:1389-1389(1974).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Plasma.  
CC -1- ALLERGEN: Causes an allergic reaction in human.  
CC -1- SIMILARITY: Belongs to the ALB/AFI/VDB family.  
CC -1- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; M73993; AAA51411.1; -  
DR EMBL; X58989; CAA41735.1; -  
DR EMBL; Y17769; CAA76847.1; -  
DR EMBL; AF542068; AANI7824.1; -  
DR HSSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot.; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
KW Polymorphism.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607 SERUM ALBUMIN.  
FT DOMAIN 25 204 ALBUMIN 1.  
FT DOMAIN 211 396 ALBUMIN 2.  
FT DOMAIN 403 594 ALBUMIN 3.  
FT METAL 27 27 COPPER (BY SIMILARITY).  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
FT VARIANT 214 214  
FT CONFLICT 302 302  
FT CONFLICT 304 305  
FT CONFLICT 324 324  
FT CONFLICT 394 395  
FT CONFLICT 437 437  
FT CONFLICT 493 494  
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
A -> T.  
C -> K (IN REF. 6).  
KP -> PC (IN REF. 6).  
N -> D (IN REF. 6).  
ST -> TS (IN REF. 6).  
K -> R (IN REF. 12).  
SE -> RS (IN REF. 6).  
Query Match 82.9%; Score 58; DB 1; Length 607;  
Best Local Similarity 84.6%; Pred.No. 0.005;  
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VASLRETYGEMAD 14  
DB 101 VASLRETYGDMAD 113  
RESULT 6  
ALBU FELCA  
ID ALBU FELCA STANDARD; PRT; 608 AA.  
AC P49064;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Fel d 2).  
GN ALB.  
OS Felis silvestris catus (Cat).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_taxid=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96194824; PubMed=8647469;

```

FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 147 192 BY SIMILARITY.
FT DISULFID 191 200 BY SIMILARITY.
FT DISULFID 223 269 BY SIMILARITY.
FT DISULFID 268 276 BY SIMILARITY.
FT DISULFID 288 302 BY SIMILARITY.
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 383 392 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 88.6%; Score 62; DB 1; Length 607;
Best Local Similarity 92.3%; Pred. No. 0.00089;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
Db 101 VATLRETYGEMAD 113

RESULT 4
ALBU_HORSE
ID ALBU_HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
PP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X74045; CAA52194.1; -

```

```

DR PIR; S34053; ABHOS.
DR HSSP; P02768; LE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Transport_prot. 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 211 396 ALBUMIN 1.
FT DOMAIN 403 594 ALBUMIN 2.
FT METAL 27 27 ALBUMIN 3.
FT DISULFID 77 86 COPPER (BY SIMILARITY).
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6B830A1B90C5 CRC64;

Query Match 87.1%; Score 61; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. No. 0.0014;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 5
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
PP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
PP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
PP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]

```

Electrophoresis 15:1459-1465(1994).  
[14]  
RN DISULFIDE BONDS.  
RP Saber M.A., Stockbauer P., Moravsek L., Meloun B.;  
RA "Disulfide bonds in human serum albumin.";  
RT Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RN BILIRUBIN-BINDING SITE.  
RP MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RN VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RN VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M., Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M., Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese albumins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RN VARIANTS MANAUS; OKAZAKI; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RN DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of human serum albumin whose precursor has an aberrant signal peptidase cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J., Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RN VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M., Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23; KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E., Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).

[24]  
RN VARIANT CASEBROOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum albumin: albumin Casebrook (494 Asp-->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H., Rochu D., Porta F.;  
RT "Two albumins with identical electrophoretic mobility are produced Query Match 100.0%; Score 70; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 2.8e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TVATLRETYGEMAD 14  
DB 100 TVATLRETYGEMAD 113  
RESULT 3  
ALBU SHEEP  
ID ALBU SHEEP STANDARD; PRT; 607 AA.  
AC P14839;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.  
OC NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver.  
RX MEDLINE=90098888; PubMed=2602160;  
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;  
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";  
RL Nucleic Acids Res. 17:10495-10495(1989).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X17055; CAA34903.1; -.  
DR F01; S06936; ABSHS.  
DR HSP; P02768; IY7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
KW SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24 BY SIMILARITY.

FT METAL 19 19 COPPER (BY SIMILARITY).  
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).  
 FT DISULFID 69 78 BY SIMILARITY.  
 FT DISULFID 91 107 BY SIMILARITY.  
 FT DISULFID 106 117 BY SIMILARITY.  
 FT DISULFID 140 185 BY SIMILARITY.  
 FT DISULFID 184 193 BY SIMILARITY.  
 FT DISULFID 216 262 BY SIMILARITY.  
 FT DISULFID 261 269 BY SIMILARITY.  
 FT DISULFID 281 295 BY SIMILARITY.  
 FT DISULFID 294 305 BY SIMILARITY.  
 FT DISULFID 332 377 BY SIMILARITY.  
 FT DISULFID 376 385 BY SIMILARITY.  
 FT DISULFID 408 454 BY SIMILARITY.  
 FT DISULFID 453 464 BY SIMILARITY.  
 FT DISULFID 477 493 BY SIMILARITY.  
 FT DISULFID 492 503 BY SIMILARITY.  
 FT DISULFID 530 575 BY SIMILARITY.  
 FT DISULFID 574 583 BY SIMILARITY.  
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 70; DB 1; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14  
 Db 92 TVATLRETYGEMAD 105

RESULT 2  
 ALBU HUMAN STANDARD; PRT; 609 AA.  
 AC P02768; Q95574; Q13140; Q9P117; Q9UHS3; Q9UJZ0;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_taxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=3009475;  
 RX MEDLINE=86196112; PubMed=3009475;  
 RA Minghetti P.P., Ruffner D.B., Kuang W.J., Dennison O.E., Hawkins J.W.,  
 RA Beattie W.G., Dugaiczky A.;  
 RT "Molecular structure of the human albumin gene is revealed by  
 RT nucleotide sequence within q11-22 of chromosome 4.";  
 RL J. Biol. Chem. 261:6747-6757(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
 RX MEDLINE=82081982; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.;  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
 RX MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.;  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).

RC TISSUE=Fetal liver;  
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
 RT "Functional prediction of the coding sequences of 121 new genes  
 RT deduced by analysis of cDNA clones from human fetal liver."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
 RA Huang M.C., Wu H.T.;  
 RT "The cDNA sequences of human serum albumin."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver, and Skeletal muscle;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold Z.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haileh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Young A.C., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 25-609.  
 RX MEDLINE=76187907; PubMed=1225573;  
 RA Meloun B., Moravsek L., Kostka V.;  
 RT "Complete amino acid sequence of human serum albumin."  
 RL FEBS Lett. 58:134-137(1975).  
 RN [9]  
 RP SEQUENCE OF 25-609.  
 RA Brown J.R., Shockley P., Behrens P.Q.;  
 RL (In) Bing D.H. (eds.);  
 RT The chemistry and physiology of the human plasma proteins, pp.23-40,  
 RL Pergamon Press, New York (1979).  
 RN [10]  
 RP SEQUENCE OF 1-455 FROM N.A.  
 RC TISSUE=Liver;  
 RA Menaya J., Parrilla R., Ayuso M.S.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [11]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=86140099; PubMed=2419329;  
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
 RT "The human albumin gene. Characterization of the 5' and 3' flanking  
 RT regions and the polymorphic gene transcripts."  
 RL J. Biol. Chem. 261:3244-3251(1986).  
 RN [12]  
 RP SEQUENCE OF 222-229.  
 RX MEDLINE=76257808; PubMed=955075;  
 RA Walker J.E.;  
 RT "Lysine residue 199 of human serum albumin is modified by  
 RT acetylsalicylic acid."  
 RL FEBS Lett. 66:173-175(1976).  
 RN [13]  
 RP SEQUENCE OF 25-44 AND 480-499.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 1.02124 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_76\_89

Perfect score: 70

Sequence: 1 TVATLRETYCEWAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 70    | 100.0       | 600    | 1     | ALBU_MACMU  |
| 2          | 70    | 100.0       | 609    | 1     | ALBU_HUMAN  |
| 3          | 62    | 88.6        | 607    | 1     | ALBU_SHEEP  |
| 4          | 61    | 87.1        | 607    | 1     | ALBU_HORSE  |
| 5          | 58    | 82.9        | 607    | 1     | ALBU_BOVIN  |
| 6          | 57    | 81.4        | 608    | 1     | ALBU_FELCA  |
| 7          | 54    | 77.1        | 608    | 1     | ALBU_CANFA  |
| 8          | 46    | 65.7        | 608    | 1     | ALBU_MOUSE  |
| 9          | 43    | 61.4        | 608    | 1     | ALBU_RABIT  |
| 10         | 42    | 60.0        | 256    | 1     | PHMX_MOUSE  |
| 11         | 42    | 60.0        | 605    | 1     | ALBU_PIG    |
| 12         | 42    | 60.0        | 608    | 1     | ALBU_RAT    |
| 13         | 41    | 58.6        | 615    | 1     | ALBU_CHICK  |
| 14         | 41    | 58.6        | 1034   | 1     | GCSP_YEAST  |
| 15         | 40    | 57.1        | 409    | 1     | ODCV_HA8IN  |
| 16         | 40    | 57.1        | 563    | 1     | PK66_PLAKU  |
| 17         | 39    | 55.7        | 419    | 1     | LWA_HYDEC   |
| 18         | 38    | 54.3        | 420    | 1     | YTF3_RHIT   |
| 19         | 38    | 54.3        | 780    | 1     | PPSA_DEIRA  |
| 20         | 37    | 52.9        | 245    | 1     | NDDB_CAVPO  |
| 21         | 37    | 52.9        | 312    | 1     | OGGI_MERTH  |
| 22         | 37    | 52.9        | 488    | 1     | IMDH_BACSU  |
| 23         | 37    | 52.9        | 759    | 1     | TOP3_CABEL  |
| 24         | 37    | 52.9        | 1261   | 1     | APAF_BRARE  |
| 25         | 36    | 51.4        | 283    | 1     | Y43F_MYCPN  |
| 26         | 36    | 51.4        | 352    | 1     | RFC5_NEUCR  |
| 27         | 36    | 51.4        | 433    | 1     | APB_CHICK   |
| 28         | 36    | 51.4        | 562    | 1     | EZRA_BACSU  |
| 29         | 36    | 51.4        | 614    | 1     | DNAX_LACAC  |
| 30         | 36    | 51.4        | 736    | 1     | VF4_ROTFC   |
| 31         | 36    | 51.4        | 1737   | 1     | CO4_FATP    |
| 32         | 36    | 51.4        | 3988   | 1     | POLG_BVDVN  |
| 33         | 35.5  | 50.7        | 3164   | 1     | TEGU_HSV11  |

#### RESULT 1

| ID | ALBU_MACMU   | STANDARD;                         | PRT; | 600 AA.        |
|----|--|-----------------------------------|------|----------------|
| AC | Q28522;  |                                   |      |                |
| DT | 01-NOV-1997  | (Rel. 35, Created)                |      |                |
| DT | 01-NOV-1997  | (Rel. 35, Last sequence update)   |      |                |
| DT | 28-FEB-2003  | (Rel. 41, Last annotation update) |      |                |
| DE | Serum albumin precursor (Fragment).  |                                   |      |                |
| GN | ALB.   |                                   |      |                |
| OS | Macaca mulatta (Rhesus macaque).   |                                   |      |                |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |                                   |      |                |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  |                                   |      |                |
| OC | Cercopithecoidea; Macaca.  |                                   |      |                |
| OX | NCBI_TaxID=9544;   |                                   |      |                |
| RN | [1]  |                                   |      |                |
| RP | SEQUENCE FROM N.A.   |                                   |      |                |
| RX | MEDLINE=93211971; PubMed=8460152;  |                                   |      |                |
| RA | Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,   |                                   |      |                |
| RA | Dwulet J., Putnam F.W.;  |                                   |      |                |
| RT | "cDNA and protein sequence of polymorphic macaque albumins that differ   |                                   |      |                |
| RT | in bilirubin binding,"   |                                   |      |                |
| RL | Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  |                                   |      |                |
| CC | 1- FUNCTION: Serum albumin, the main protein of plasma, has a good   |                                   |      |                |
| CC | binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  |                                   |      |                |
| CC | hormones, bilirubin and drugs. Its main function is the regulation   |                                   |      |                |
| CC | of the colloidal osmotic pressure of blood.  |                                   |      |                |
| CC | 1- TISSUE SPECIFICITY: Plasma.   |                                   |      |                |
| CC | 1- SUBCELLULAR LOCATION: Secreted.   |                                   |      |                |
| CC | 1- SIMILARITY: Belongs to the ALB/APP/VDB family.  |                                   |      |                |
| CC | 1- SIMILARITY: Contains 3 albumin domains.   |                                   |      |                |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration   |                                   |      |                |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |                                   |      |                |
| CC | the European Bioinformatics Institute. There are no restrictions on its  |                                   |      |                |
| CC | use by non-profit institutions as long as its content is in no way   |                                   |      |                |
| CC | modified and this statement is not removed. Usage by and for commercial  |                                   |      |                |
| CC | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |                                   |      |                |
| CC | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |                                   |      |                |
| CC | EMBL; M90463; AAA36906.1; -  |                                   |      |                |
| DR | PIR; A47391; A47391.   |                                   |      |                |
| DR | HSP; P02768; 1E7B.   |                                   |      |                |
| DR | InterPro; IPR000264; Serum albumin.  |                                   |      |                |
| DR | Pfam; PF00273; transport_prot; 3.  |                                   |      |                |
| DR | PRINTS; PR00802; SERUMALBUMIN.   |                                   |      |                |
| DR | ProDom; PD002486; Serum albumin; 1.  |                                   |      |                |
| DR | SMART; SM00103; ALBUMIN; 3.  |                                   |      |                |
| DR | PROSITE; PS00212; ALBUMIN; 3.  |                                   |      |                |
| KW | Metal-binding; Lipid-binding; Repeat; Signal; Copper.  |                                   |      |                |
| FT | NON_TER  | 1                                 |      |                |
| FT | SIGNAL   | <1                                | 10   | BY SIMILARITY. |
| FT | PROPEP   | 11                                | 16   | BY SIMILARITY. |
| FT | CHAIN  | 17                                | 600  | SERUM ALBUMIN. |
| FT | DOMAIN   | 17                                | 197  | ALBUMIN 1.     |
| FT | DOMAIN   | 204                               | 389  | ALBUMIN 2.     |
| FT | DOMAIN   | 396                               | 587  | ALBUMIN 3.     |

Q09921 schizosacch  
P76506 escherichia  
P43262 shigella fl  
O28550 archaeoglob  
Q28606 chlamydia p  
P19730 pseudomonas  
P23359 saccharomyc  
P36559 escherichia  
Q83q00 shigella fl  
O26802 methanobact  
Q9f314 streptomyces  
O29632 archaeoglob

#### ALIGNMENTS

A;Cross-references: EMBL:Z47815; NID:G642280; PIDN:CAA87810.1; PID:G642281; MIPS:YMR189C  
 R;Sinclair, D.A.; Dawes, I.W.  
 submitted to the EMBL Data Library, February 1995  
 A;Description: Cloning and expression of glycine decarboxylase from *Saccharomyces cerevisiae*  
 A;Reference number: S59809  
 A;Accession: S59810  
 A;Molecule type: DNA  
 A;Residues: 1-1034 <SIN>  
 A;Cross-references: EMBL:U20641; NID:G676869; PIDN:AA818933.1; PID:G676871  
 R;Sinclair, D.A.; Hong, S.P.; Dawes, I.W.  
 Mol. Microbiol. 19, 611-623, 1996  
 A;Title: Specific induction by glycine of the gene for the P-subunit of glycine decarboxylase  
 A;Reference number: S70896; MUID:96228709; PMID:8830251  
 A;Accession: S70896  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 1-77, 'V', 79-121, 'VS', 124-247, 'L', 249-381, 'T', 383-517, 'H', 519-1034 <SIW>  
 A;Cross-references: EMBL:U20641  
 C;Genetics:  
 A;Gene: SGD:GCV2; GCV2; GSD2  
 A;Cross-references: SGD:S0004801  
 A;Map position: 13R  
 C;Keywords: phosphoprotein; pyridoxal phosphate; transferase  
 F;773/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 58.8%; Score 41; DB 2; Length 1034;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVATLRETYGEM 12  
 ||||| :|||:  
 DB 1006 TVARLDDTYGDM 1017

RESULT 15  
 A95074  
 ABC transporter, ATP-binding protein SP0636 [imported] - *Streptococcus pneumoniae* (strain  
 C;Species: *Streptococcus pneumoniae*  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
 C;Accession: A95074  
 R;Rettstein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
 nson, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A;Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.  
 A;Reference number: A95000; MUID:21357289; PMID:11463916  
 A;Accession: A95074  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-330 <KUR>  
 A;Cross-references: GB:AE005672; PIDN:AAK74786.1; PID:G14972111; GSPDB:GN00164; TIGR:SP4  
 A;Experimental source: strain TIGR4  
 C;Genetics:  
 A;Gene: SP0636  
 C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 57.1%; Score 40; DB 2; Length 330;  
 Best Local Similarity 58.3%; Pred. No. 16;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRETYGEM 12  
 ||: |||||:  
 DB 240 TVSQLKTFGKM 251

Search completed: April 19, 2004, 12:02:23  
 Job time : 2.70637 secs

A;Cross-references: GB:V00381; NID:G63038; PIDN:CRA23680.1; PID:G63039  
R;Rosen, A.M.; Geiler, D.M.  
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977  
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.  
A;Reference number: A13451; MUID:78019943; PMID:911327  
A;Accession: A13451  
A;Molecule type: protein  
A;Residues: 19-23, 'M', 25-30 <ROS>  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thy C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-26/Domain: propeptide #status predicted <PRO>  
F;27-613/Product: serum albumin #status predicted <MAT>  
F;32-206/Domain: serum albumin repeat homology <SA1>  
F;225-398/Domain: serum albumin repeat homology <SA2>  
F;417-596/Domain: serum albumin repeat homology <SA3>  
F;30/Binding site: copper (his) #status predicted  
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397

Query Match 58.6%; Score 41; DB 1; Length 615;  
Best Local Similarity 61.5%; Pred. No. 20;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
DB 104 VEKLRDSYGAMAD 116

RESULT 13  
G87415  
glycyl-tRNA synthetase, beta subunit [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C;Accession: G87415  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, C.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc B. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.N. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: G87415  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-692 <STO>  
A;Cross-references: GB:AF005673; NID:gl3422689; PIDN:AAK23323.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CCI342  
C;Superfamily: glycine-tRNA ligase beta chain

Query Match 58.6%; Score 41; DB 2; Length 692;  
Best Local Similarity 50.0%; Pred. No. 23;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
| | | | | | | | | | | | | | | |  
DB 673 TLATVRDMGQVAD 686

RESULT 14  
S50917  
aminomethyltransferase (EC 2.1.1.2.10) - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: glycyl cleavage system protein P; protein YM9646.01; protein YMF189  
C;Species: Saccharomyces cerevisiae  
C;Date: 10-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 19-Apr-2002  
C;Accession: S50917; S59910; S70896  
R;Pearson, D. J. Bowman, S.  
submitted to the EMBL Data Library, January 1995  
A;Reference number: S50917  
A;Accession: S50917  
A;Molecule type: DNA  
A;Residues: 1-1034 <PEA>

F:123-236/Domain: serum albumin repeat homology <SA2>  
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>  
Query Match 65.7%; Score 46; DB 2; Length 453;  
Best Local Similarity 61.5%; Pred. No. 1.7;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14  
: ||||| |||  
DB 3 IPNLRENYGELAD 15

## RESULT 8

AH2433  
hypothetical protein all5024 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AH2433  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759940  
A:Accession: AH2433  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-252 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA076723.1; PID:gl17134162; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all5024

Query Match 60.0%; Score 42; DB 2; Length 252;  
Best Local Similarity 64.3%; Pred. No. 4.9;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14  
: ||||| |||  
DB 176 TIATLRSTAHMWD 189

## RESULT 9

ABPGS  
serum albumin precursor - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S01382; A61006  
R:Wainstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 5045, 1988  
A:Title: Nucleotide sequence of porcine liver albumin.  
A:Reference number: S01382; MUID:89016582; PMID:3174440  
A:Accession: S01382  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-605 <WEI>  
A:Cross-references: EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g933798  
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral  
A:Reference number: A61006; MUID:89269769; PMID:2728927  
A:Accession: A61006  
A:Molecule type: protein  
A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>  
A:Experimental source: dental enamel  
A:Note: albumin and other serum proteins are also found in bone  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membra  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-22/Domain: propeptide #status predicted <PRO>  
F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>  
F:218-391/Domain: serum albumin repeat homology <SA2>  
F:410-589/Domain: serum albumin repeat homology <SA3>  
F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390,  
F:76/Binding site: bilirubin (lys) #status predicted  
Query Match 60.0%; Score 42; DB 1; Length 605;  
Best Local Similarity 53.8%; Pred. No. 13;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14  
: ||||| |||  
DB 99 IPSLREHYGLAD 111

## RESULT 10

ABRTS  
serum albumin precursor - rat  
N:Alternate names: preproalbumin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-May-1979 #sequence\_revision 31-May-1979 #text\_change 22-Jun-1999  
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233  
R:Sargent, T.D.; Yang, M.; Bonner, J.  
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981  
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
A:Reference number: A93872; MUID:81223722; PMID:7017712  
A:Accession: A93872  
A:Molecule type: mRNA  
A:Residues: 1-608 <SAR>  
A:Cross-references: GB:V01222; GB:J00698; NID:g55627; PIDN:CAA24532.1; PID:g55628  
R:Stauss, A.W.; Bennett, C.D.; Donchue, A.M.; Rodkey, J.A.; Alberts, A.W.  
J. Biol. Chem. 252, 6846-6855, 1977  
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Anal.  
A:Reference number: A92211; MUID:77249657; PMID:893447  
A:Note: cleavages during protein maturation  
A:Accession: A92211  
A:Molecule type: protein  
A:Residues: 1-38 <STR>  
R:Isemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleav  
A:Reference number: A91946; MUID:78109429; PMID:564345  
A:Accession: A91946  
A:Molecule type: protein  
A:Residues: 25-222 <ISI>  
R:Isemura, S.; Ikenaka, T.  
J. Biochem. 79, 1183-1196, 1976  
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino  
A:Reference number: A91940; MUID:76260153; PMID:956149  
A:Accession: A91940  
A:Molecule type: protein  
A:Residues: 223-288; 572-608 <IS2>  
A:Note: 262-Leu was also found  
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A:Reference number: A30758; MUID:79001617; PMID:80265  
A:Contents: annotation; copper binding  
J. Immunol. 143, 1680-1684, 1989  
R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
A:Title: Structures of histamine-releasing peptides formed by the action of acid protea  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: C45800  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <CAR>  
R:Heard, J.  
Mol. Cell. Biol. 7, 2425-2434, 1987  
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improve  
A:Reference number: I57621; MUID:87286876; PMID:3475566  
A:Accession: I57621  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A;Reference number: A91258; MUID:80024278; PMID:488109  
A;Accession: A91258  
A;Molecule type: protein  
A;Residues: 1-32 <MAG>  
R;Hsieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A;Title: Electrophoretic transfer of protein from an analytical isoelectrofocusing gel to a nitrocellulose membrane.  
A;Reference number: A60808; MUID:88267456; PMID:3389500  
A;Accession: B60808  
A;Molecule type: protein  
A;Residues: 25-41 <HSI>  
R;Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin.  
A;Reference number: S10780; MUID:90336641; PMID:2379503  
A;Accession: S10780  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>  
R;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases.  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: A45800  
A;Molecule type: protein  
A;Residues: 163-172 <CAR>  
R;Caraway, R.E.; Mitra, S.P.; Cochran, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A;Title: Structure of a biologically active neurotensin-related peptide obtained from peptidomimetic synthesis.  
A;Reference number: A26693; MUID:87194805; PMID:2437111  
A;Accession: A26693  
A;Molecule type: protein  
A;Residues: 163-172, 'L' <CA2>  
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A;Title: Sequence of residues 400-403 of bovine serum albumin.  
A;Reference number: A30309; MUID:82023364; PMID:7283978  
A;Accession: A30309  
A;Molecule type: protein  
A;Residues: 402-433 <REE>  
R;Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A;Title: Structure of bovine serum albumin.  
A;Reference number: A91458  
A;Accession: A91458  
A;Molecule type: protein  
A;Residues: 190-195 <BR2>  
R;Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A;Reference number: A91457  
A;Contents: annotation; disulfide bonds  
R;Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC 3.4.11.18).  
A;Reference number: S55232; MUID:95031935; PMID:7945219  
A;Accession: S55232  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 529-536; 569-572 <WER>  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; copper binding; duplication; plasma  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-607/Product: serum albumin #status experimental <MPT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>

F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,  
Query Match 82.9%; Score 58; DB 1; Length 607;  
Best Local Similarity 84.6%; Pred. No. 0.015;  
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VATLRETYGEMAD 14  
DB 101 VASLRDXYGEMAD 113  
RESULT 6  
S57632  
serum albumin precursor - cat  
C;Species: Felis silvestris catus (domestic cat)  
C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C;Accession: JC4660; S57632  
R;Hilger, C.; Grigioni, F.; Hentges, F.  
Gene 169, 295-296, 1996  
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A;Reference number: JC4660; MUID:96194824; PMID:8647469  
A;Accession: JC4660  
A;Molecule type: mRNA  
A;Residues: 1-608 <HI2>  
A;Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
A;Experimental source: liver  
C;Comment: This protein is the major protein component in plasma. It functions as a multimeric protein.  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: liver; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PR>  
F;25-608/Product: serum albumin #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
Query Match 81.4%; Score 57; DB 2; Length 608;  
Best Local Similarity 78.6%; Pred. No. 0.023;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TVATLRETYGEMAD 14  
DB 100 TVASLRDXYGEMAD 113  
RESULT 7  
A05139  
serum albumin - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 05-Jun-1987 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000  
C;Accession: A05139; I48638  
R;Minghetti, P.F.; Law, S.W.; Dugalczyk, A.  
Mol. Biol. Evol. 2, 347-358, 1985  
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes.  
A;Reference number: A93055; MUID:88216123; PMID:2452956  
A;Accession: A05139  
A;Molecule type: mRNA  
A;Residues: 1-418 <MIN>  
A;Cross-references: GB:ML6111; NID:G191764; PIDN:AAA37190.1; PID:G191765  
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.  
Gene 88, 181-186, 1990  
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the human genome.  
A;Reference number: I48638  
A;Accession: I48638  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 379-453 <BOC>  
A;Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>

A;Contents: annotation; binding sites  
R;Harper, M.E.; Dugaiczak, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
A;Reference number: A90028; MUID:8327982; PMID:6192711  
A;Contents: annotation; gene position  
R;Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
A;Reference number: A46755; MUID:76257808; PMID:955075  
A;Contents: annotation  
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate  
A;Reference number: A56294; MUID:92183881; PMID:1544460  
A;Contents: annotation  
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p  
ataze activity  
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
C;Comment: A large number of variants of human serum albumin have been described.  
C;Genetics:  
A;Gene: GDB:ALB  
A;Cross-references: GDB:118990; OMIM:103600  
A;Map Position: 4q11-4q13  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-609/Product: serum albumin #status experimental <MPt>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;166-174/Product: kinetensin #status experimental <KIP>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental  
Query Match 100.0%; Score 70; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 9.1e-05; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;  
QY 1 TVATLRETYGEMAD 14  
|||||  
DB 100 TVATLRETYGEMAD 113

RESULT 3  
ABSHS  
serum albumin precursor - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: S06936  
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A;Reference number: S06936; MUID:9009888; PMID:2602160  
A;Accession: S06936  
A;Molecule type: mRNA  
A;Residues: 1-607 <PRO>  
A;Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membra  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRO>  
F;25-607/Product: serum albumin #status predicted <MAT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted

F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,  
F;263/Binding site: bilirubin (Lys) #status predicted  
Query Match 88.6%; Score 62; DB 1; Length 607;  
Best Local Similarity 92.3%; Pred. No. 0.0027;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VATLRETYGEMAD 14  
|||||  
DB 101 VATLRETYGEMAD 113

RESULT 4  
ABHOS  
serum albumin precursor - horse  
C;Species: Equus caballus (domestic horse)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: S34053  
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
Eur. J. Biochem. 215, 205-212, 1993  
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A;Reference number: S34053; MUID:93345495; PMID:8344282  
A;Accession: S34053  
A;Molecule type: mRNA  
A;Residues: 1-607 <HOA>  
A;Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membra  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRO>  
F;25-607/Product: serum albumin #status predicted <MAT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,  
F;263/Binding site: bilirubin (Lys) #status predicted  
Query Match 87.1%; Score 61; DB 1; Length 607;  
Best Local Similarity 85.7%; Pred. No. 0.0041;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TVATLRETYGEMAD 14  
|||||  
DB 100 TVATLRETYGEMAD 113

RESULT 5  
ABEOS  
serum albumin precursor [validated] - bovine  
N;Alternate names: 67K protein; preproalbumin  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 24-Apr-1984 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000  
C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A:  
R;Holowachuk, E.W.; Stølenborg, J.K.; Reed, R.G.; Peters Jr., T.  
submitted to the EMBL Data Library, August 1991  
A;Description: Bovine serum albumin: cDNA sequence and expression.  
A;Reference number: A38885  
A;Accession: A38885  
A;Molecule type: mRNA  
A;Residues: 1-607 <HOA>  
A;Cross-references: EMBL:M73215  
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
A;Title: Rapid confirmation and revision of the primary structure of bovine serum album  
A;Reference number: A36401; MUID:91083649; PMID:2260975  
A;Accession: A36401  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>  
R;Macgillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979

A/Accession: A93936  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-609 <DUG>  
A/Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590  
R/Uranio, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and A/Reference number: I39427; MUID:86140099; PMID:2419329  
A/Accession: I39427  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-26 <URA>  
A/Cross-references: GB:M13075; NID:G178330; PIDN:AA51688.1; PID:G553173  
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A/Reference number: I59286; MUID:94181575; PMID:8134387  
A/Accession: I59286  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 282-290, 'KSRFDLO' <WAT>  
A/Cross-references: GB:S69192; NID:G546032; PIDN:RAB30282.1; PID:G546033  
A/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl- A/Reference number: I59313; MUID:94294404; PMID:8022807  
A/Accession: I59313  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 589-590, 'ALPRVKNLLQVKLP' <MAD>  
A/Cross-references: GB:S70759; NID:G547231; PIDN:RAB31177.1; PID:G547232  
A/Note: this frame-shift variant is designated albumin Bazzano; four additional variants submitted to the EMBL Data Library, March 1995  
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.  
A/Reference number: G08252  
A/Accession: G01747  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-455 <MEN>  
A/Cross-references: EMBL:U22963; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2 A/Reference number: S55314; MUID:95275251; PMID:7755581  
A/Accession: S55314  
A/Molecule type: protein  
A/Residues: 19-27 <LED>  
R/Meloun, B.; Moravek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A/Title: Complete amino acid sequence of human serum albumin.  
A/Reference number: A91420; MUID:76187907; PMID:1225573  
A/Accession: A91420  
A/Molecule type: protein  
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395- R/Koehn, U.; Spittler, G.; Tripiet, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from u A/Reference number: S06422  
A/Note: this paper is in German, with an English abstract  
A/Accession: S06422  
A/Molecule type: protein  
A/Residues: 25-48 <ROB>  
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A/Title: Mass spectrometric identification of modifications to human serum albumin treat A/Reference number: S36882; MUID:93384321; PMID:8373198  
A/Accession: S36882  
A/Molecule type: protein  
A/Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <PIN>  
R/Kausler, E.; Spittler, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A/Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A/Reference number: S17599; MUID:92126241; PMID:1772598  
A/Accession: S17599  
A/Molecule type: protein  
A/Residues: 25-54, 154-357, 431-447 <KAU>  
A/Note: 49-Leu was also found  
R/Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1694, 1989  
A/Title: Structures of histamine-releasing peptides formed by the action of acid protease A/Reference number: A45800; MUID:89341406; PMID:2474609  
A/Accession: A45800  
A/Molecule type: protein  
A/Residues: 166-173, 'L' <MOG>  
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; P Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-ti A/Reference number: A03239; MUID:86242180; PMID:3087352  
A/Accession: A03239  
A/Molecule type: protein  
A/Residues: 76-83, 'K', 85-106 <GAL3>  
A/Note: this variant is designated albumin Torino  
R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Temi, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A/Title: The structural characterization and bilirubin-binding properties of albumin He A/Reference number: S33298; MUID:93292504; PMID:8513793  
A/Accession: S33298  
A/Molecule type: protein  
A/Residues: 255-263, 'E', 265-281 <MIN1>  
A/Note: this variant is designated albumin Herborn  
R/Minchiotti, L.; Galliano, M.; Stoppani, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta Biochim. Biophys. Acta 1119, 232-238, 1992  
A/Title: Two albumins with identical electrophoretic mobility are produced by diff A/Reference number: S21078; MUID:92190239; PMID:1347703  
A/Accession: S21078  
A/Molecule type: protein  
A/Residues: 354-356, 'K', 358-378 <MIN2>  
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported R/He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A/Title: Atomic structure and chemistry of human serum albumin.  
A/Reference number: A46756; MUID:92334427; PMID:1630489  
A/Contents: annotation; X-ray crystallography, 2.8 angstroms  
R/Brown, J.R.; Shockley, P.; Behrens, F.Q.  
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40 A/Reference number: A94442  
A/Contents: annotation; three-dimensional structure and disulfide bonds  
R/Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A/Title: Disulfide bonds in human serum albumin.  
A/Reference number: A90930  
A/Contents: annotation; disulfide bonds  
R/Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A/Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding A/Reference number: A90299; MUID:78186630; PMID:656055  
A/Contents: annotation; bilirubin-binding site  
R/Peters, T.; Reed, R.G.  
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20 A/Title: Serum albumin: conformation and active sites.  
A/Reference number: A94408

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.70637 seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_copy\_76\_89

Perfect score: 70  
Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: Piri:\*

2: Piri:\*

3: Piri:\*

4: Piri:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 70    | 100.0       | 600    | 2 A47391 | serum albumin prec  |
| 2          | 70    | 100.0       | 609    | 1 ABHUS  | serum albumin prec  |
| 3          | 62    | 88.6        | 607    | 1 ABHSH  | serum albumin prec  |
| 4          | 61    | 87.1        | 607    | 1 ABHOS  | serum albumin prec  |
| 5          | 58    | 82.9        | 607    | 1 ABHOS  | serum albumin prec  |
| 6          | 57    | 81.4        | 608    | 2 S57632 | serum albumin prec  |
| 7          | 46    | 65.7        | 483    | 2 A05139 | serum albumin prec  |
| 8          | 42    | 60.0        | 282    | 2 AB2433 | hypothetical prote  |
| 9          | 42    | 60.0        | 605    | 1 ABPSS  | serum albumin prec  |
| 10         | 42    | 60.0        | 608    | 1 ABRTS  | serum albumin prec  |
| 11         | 41    | 58.6        | 354    | 2 H89009 | protein R08F11.4 [  |
| 12         | 41    | 58.6        | 615    | 1 ABCHS  | serum albumin prec  |
| 13         | 41    | 58.6        | 692    | 2 G87415 | glycyl-tRNA synthet |
| 14         | 41    | 58.6        | 1034   | 2 S50917 | aminomethyltransfe  |
| 15         | 40    | 57.1        | 330    | 2 A95074 | ABC transporter, A  |
| 16         | 40    | 57.1        | 330    | 2 E97941 | hypothetical prote  |
| 17         | 40    | 57.1        | 409    | 2 D64135 | dihydrodiploamide s |
| 18         | 40    | 57.1        | 563    | 2 A39238 | 66K merozoite surf  |
| 19         | 40    | 57.1        | 585    | 2 B75265 | probable acid-CoA   |
| 20         | 39.5  | 56.4        | 202    | 2 T31452 | probable 4-vinyl r  |
| 21         | 39    | 55.7        | 157    | 2 T22606 | hypothetical prote  |
| 22         | 39    | 55.7        | 179    | 2 C86853 | hypothetical prote  |
| 23         | 38    | 54.3        | 251    | 2 AF0805 | Vacu lipoprotein p  |
| 24         | 38    | 54.3        | 780    | 2 D75361 | phosphoenolpyruvat  |
| 25         | 38    | 54.3        | 867    | 2 F81086 | aminopeptidase N    |
| 26         | 37.5  | 53.6        | 1350   | 2 T42697 | hypothetical prote  |
| 27         | 37    | 52.9        | 169    | 2 AE3257 | shikimate kinase    |
| 28         | 37    | 52.9        | 228    | 2 D96945 | response regulator  |
| 29         | 37    | 52.9        | 231    | 2 T46373 | hypothetical prote  |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 37   | 52.9 | 312  | 2 E69045 | 8-oxoguanine DNA g |
| 31 | 37   | 52.9 | 332  | 2 H83924 | tryptophanyl-tRNA  |
| 32 | 37   | 52.9 | 488  | 1 DEBSMP | IMP dehydrogenase  |
| 33 | 37   | 52.9 | 759  | 2 T43031 | DNA topoisomerase  |
| 34 | 37   | 52.9 | 1172 | 2 AD2310 | hypothetical prote |
| 35 | 37   | 52.9 | 1278 | 2 T27925 | hypothetical prote |
| 36 | 36.5 | 52.1 | 378  | 2 T39790 | hypothetical prote |
| 37 | 36   | 51.4 | 69   | 2 D95250 | hypothetical prote |
| 38 | 36   | 51.4 | 89   | 2 A69513 | probable dimethyl  |
| 39 | 36   | 51.4 | 145  | 2 C85896 | hypothetical prote |
| 40 | 36   | 51.4 | 157  | 2 B83496 | hypothetical prote |
| 41 | 36   | 51.4 | 205  | 2 AH0403 | anaerobic dimethyl |
| 42 | 36   | 51.4 | 209  | 2 G91051 | hypothetical prote |
| 43 | 36   | 51.4 | 256  | 2 D64637 | amino acid ABC tra |
| 44 | 36   | 51.4 | 257  | 2 E71877 | probable amino aci |
| 45 | 36   | 51.4 | 287  | 2 S73529 | M6439 homolog 809_ |

ALIGNMENTS

RESULT 1

A47391 serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bili  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homoygote  
A:Accession: A47391  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)  
F:21-194/Domain: serum albumin repeat homology <SA1>  
F:213-386/Domain: serum albumin repeat homology <SA2>  
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 70; DB 2; Length 600;  
Best Local Similarity 100.0%; Pred. No. 8.9e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TVATLRETYGEMAD 14

Db 92 TVATLRETYGEMAD 105

RESULT 2

ABHUS serum albumin precursor [validated] - human  
N:Alternate names: preproalbumin  
N:Contains: kinetensin  
C:Species: Homo sapiens (man)  
C:Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text\_change 17-Mar-2000  
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3  
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col  
A:Reference number: A93743; MUID:82081882; PMID:6171778  
A:Accession: A93743  
A:Molecule type: mRNA  
A:Residues: 1-419; 'K', 421-609 <LAW>  
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2  
R:Dugaiczky, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A:Reference number: A93936; MUID:82105994; PMID:6275391



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

```

```

Query Match      100.0%; Score 38; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 7.5; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

```

RESULT 15

```

US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: P2556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

```

```

Query Match      100.0%; Score 38; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 8.1; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      120 VADESAEN 127

```

Search completed: April 19, 2004, 12:05:19  
Job time : 2.22622 secs

```

; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-04075-3
;
; Query Match 100.0%; Score 38; DB 5; Length 609;
; Best Local Similarity 100.0%; Pred. No. 7.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 1 VADESAEN 8
Db 78 VADESAEN 85
;
; RESULT 13
; US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; INFORMATION FOR SEQ ID NO: 2:
; Query Match 100.0%; Score 38; DB 2; Length 610;
; Best Local Similarity 100.0%; Pred. No. 7.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 1 VADESAEN 8
Db 78 VADESAEN 85
;
; RESULT 14
; US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:

```

```

; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; Query Match 100.0%; Score 38; DB 2; Length 610;
; Best Local Similarity 100.0%; Pred. No. 7.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 1 VADESAEN 8
Db 78 VADESAEN 85
;
; RESULT 14
; US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:

```

APPLICANT: Bart, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschopp, Juerg F.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DiGiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9109Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 100.0%; Score 38; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
Db 78 VADESAEN 85

## RESULT 9

US-08-897-956A-2  
Sequence 2, Application US/08897956A  
Patent No. 6423512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 609  
TYPE: PPT  
ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 100.0%; Score 38; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
Db 78 VADESAEN 85

RESULT 10  
US-10-153-064-7  
Sequence 7, Application US/10153064  
Patent No. 6663485  
GENERAL INFORMATION:  
APPLICANT: Bell et al.  
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
FILE REFERENCE: PFS56  
CURRENT APPLICATION NUMBER: US/10/153,064  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 60/293,212  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 7  
LENGTH: 609  
TYPE: PPT  
ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 38; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
Db 78 VADESAEN 85

## RESULT 11

US-09-976-594-977  
Sequence 977, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 977  
LENGTH: 609  
TYPE: PPT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1  
US-09-976-594-977

Query Match 100.0%; Score 38; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
Db 78 VADESAEN 85

## RESULT 12

PCT-US95-04075-3  
Sequence 3, Application PC/TUS9504075  
GENERAL INFORMATION:  
APPLICANT: AMGEN INC.

Page 3

MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-n)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"

Query Match 100.0%; Score 38; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 54 VADESAEN 61

## RESULT 2

US-08-448-196A-3  
Sequence 3, Application US/08448196A

Patent No. 5780594  
GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.

TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
TITLE OF INVENTION: RELATED PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA

STREET: MARSHALL SPACE FLIGHT CENTER

CITY: HUNTSVILLE

STATE: ALABAMA

COUNTRY: USA

ZIP: 35812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/448,196A

FILING DATE: 23-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.

REGISTRATION NUMBER: 18,757

REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 205-544-0021

TELEFAX: 205-544-0258

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 100.0%; Score 38; DB 1; Length 585;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 54 VADESAEN 61

## RESULT 3

US-08-984-176-1  
Sequence 1, Application US/08984176

Patent No. 5948609

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C

APPLICANT: HO, JOSEPH X

APPLICANT: RUKER, FLORIAN

TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT

TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176

CURRENT APPLICATION NUMBER: US/08/984,176

CURRENT FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 585

TYPE: PRT

ORGANISM: Homo sapiens

US-08-984-176-1

Query Match 100.0%; Score 38; DB 2; Length 585;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 54 VADESAEN 61

## RESULT 4

US-08-702-572-2  
Sequence 2, Application US/08702572

Patent No. 5965386

GENERAL INFORMATION:

APPLICANT: Kerry-Williams, Sean M

APPLICANT: Gilbert, Sarah C

TITLE OF INVENTION: Yeast Strains and Modified Albumins

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Centeon L.L.C.

STREET: 1020 First Avenue

CITY: King of Prussia

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19406-1310

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,572

FILING DATE: 11-NOV-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 95/23857

FILING DATE: 1-MAR-1995

APPLICATION NUMBER: GB 9404270.2

FILING DATE: 5-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Naomi Biswas

REGISTRATION NUMBER: 38,384

REFERENCE/DOCKET NUMBER: CE0114 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610/878/4294

TELEFAX: 610/878/4221

INFORMATION FOR SEQ ID NO: 2:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.26222 Seconds  
(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61  
Perfect score: 38  
Sequences: 1 VADESAEN 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCrUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 38    | 100.0       | 585    | 1  | US-08-153-799-14  |
| 2          | 38    | 100.0       | 585    | 1  | US-08-448-196A-3  |
| 3          | 38    | 100.0       | 585    | 2  | US-08-984-176-1   |
| 4          | 38    | 100.0       | 585    | 2  | US-08-702-572-2   |
| 5          | 38    | 100.0       | 585    | 3  | US-08-769-746-2   |
| 6          | 38    | 100.0       | 585    | 4  | US-10-153-064-5   |
| 7          | 38    | 100.0       | 609    | 1  | US-08-222-619-3   |
| 8          | 38    | 100.0       | 609    | 1  | US-08-433-037-4   |
| 9          | 38    | 100.0       | 609    | 4  | US-08-897-956A-2  |
| 10         | 38    | 100.0       | 609    | 4  | US-10-153-064-7   |
| 11         | 38    | 100.0       | 609    | 4  | US-09-976-594-977 |
| 12         | 38    | 100.0       | 609    | 5  | PCT-US95-04075-3  |
| 13         | 38    | 100.0       | 610    | 2  | US-08-797-689-2   |
| 14         | 38    | 100.0       | 610    | 4  | US-09-984-186-2   |
| 15         | 38    | 100.0       | 651    | 4  | US-10-153-064-133 |
| 16         | 38    | 100.0       | 652    | 4  | US-10-153-064-96  |
| 17         | 38    | 100.0       | 652    | 4  | US-10-153-064-99  |
| 18         | 38    | 100.0       | 652    | 4  | US-10-153-064-105 |
| 19         | 38    | 100.0       | 652    | 4  | US-10-153-064-132 |
| 20         | 38    | 100.0       | 653    | 4  | US-10-153-064-131 |
| 21         | 38    | 100.0       | 656    | 4  | US-10-153-064-130 |
| 22         | 38    | 100.0       | 660    | 4  | US-10-153-064-90  |
| 23         | 38    | 100.0       | 660    | 4  | US-10-153-064-93  |
| 24         | 38    | 100.0       | 668    | 4  | US-10-153-064-102 |
| 25         | 38    | 100.0       | 676    | 4  | US-10-153-064-95  |
| 26         | 38    | 100.0       | 676    | 4  | US-10-153-064-98  |
| 27         | 38    | 100.0       | 676    | 4  | US-10-153-064-104 |

Sequence 127, App  
Sequence 129, App  
Sequence 125, App  
Sequence 123, App  
Sequence 92, Appl  
Sequence 101, App  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 89, Appl  
Sequence 7, Appl  
Sequence 5, Appl  
Sequence 4, Appl  
Sequence 4714, Ap  
Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

of the colloidal osmotic pressure of blood.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- TISSUE SPECIFICITY: Plasma.  
CC -|- SIMILARITY: Belongs to the ALB/AFB/VDB family.  
CC -|- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M18350; AAA49636.1; -;  
DR PIR; A41682; ABXL68.  
DR HSP; P02768; IZ7B.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 606  
FT DOMAIN 25 205  
FT DOMAIN 212 395  
FT DOMAIN 402 593  
FT METAL 30 30  
FT METAL 80 89  
FT DISULFID 102 118  
FT DISULFID 117 128  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 340 383  
FT DISULFID 382 391  
FT DISULFID 414 460  
FT DISULFID 459 470  
FT DISULFID 483 499  
FT DISULFID 498 509  
FT DISULFID 536 581  
FT DISULFID 580 589  
SQ SEQUENCE 606 AA; 70499 MW; F8E72AA952C3A3E8 CRC64;  
Query Match 35.1%; Score 726.5; DB 1; Length 606;  
Best Local Similarity 36.9%; Pred. No. 1.2e-46;  
Matches 139; Conservative 74; Mismatches 155; Indels 9; Gaps 4;  
QY 3 HKSEVAHFRDGLGENFKALVLAFAQVLOQCPEDHVKLVNVEVTEPAKTCVADESAENC 62  
DB 30 HKHGIADMYNLLTERTFGLTALVSNLQKCSLELSKLVNINDEPAKCTGNDKTPEC 89  
QY 63 DKSLLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHK--DDNPNLPLVYRPEV 120  
DB 90 EKPIGTLFYDKLCKADPKGVNVEYSEKCCSKQDPERAQCFRAHVPFENP-----VRPKP 144  
QY 121 DVMCTAFHDNEETLKKLYELIARRHPYFAPPELLFFAKYKAAFTCCQAADKAAQLLP 180  
DB 145 EETCALPKFHPDDLLSAFIHEEARNHNPPLYPFVALLLTQQYGLVHCCEEDKDKCFAE 204  
QY 181 KLDELDEGKASSAKQLKCAQKGFGERAFKAWAVARLSQRPFAEFABEVSKLVTDLT 240  
DB 205 KWKELMKHSHSIEDKQKHFQWVNNVYPERVVKALNARVSHRYPKDPDKLAHKTBTETH 264  
QY 241 VHTCCGHDLLCEADDRADLAKYICENQDSISSKLCCKECPKLLKSHCIAEVENDMPA 300  
DB 265 FIKDCCHGDMFECMTLELSEHTCOHKDELSTKLEKCNPLPLLERTYCIIVLENDVPA 324

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLGMPLYEYARRHPDYSVVLLRLAKTYETILEKC 360  
DB 325 ELSKPITEFTEDPHVCEKYAENKS-FLEISFWQ-SQETPELSEQFLLOSAKYEVSLLNKC 382  
QY 361 CAAADPHECYAKVFDEF 377  
DB 383 CFSNPPCEYKDGADRF 399

Search completed: April 19, 2004, 11:52:45  
Job time : 30.2299 secs





RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RA Cassidy A.I., Salkild C.K., Baverstock P., Wallace J.C.;  
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE OF 1-28 FROM N.A.  
 RP MEDLINE=83161027; PubMed=6187737;  
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deesley R.G.;  
 RT "The 5' noncoding and flanking regions of the avian very low density  
 RT apolipoprotein II and serum albumin genes. Homologies with the egg  
 RT white protein genes";  
 RL J. Biol. Chem. 258:4556-4564(1983).  
 [3]  
 RN SEQUENCE OF 19-30.  
 RP MEDLINE=78019943; PubMed=911327;  
 RA Rosen A.W., Geller D.M.;  
 RT "Chicken macroosomal albumin: amino terminal sequence of chicken  
 RT proalbumin";  
 RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).  
 [4]  
 RN ALLERGENIC PROPERTIES  
 RP MEDLINE=21381307; PubMed=11488669;  
 RA Quirice S., Maranon F., Umplierrez A., de las Heras M.,  
 RA Fernandez-Caldas E., Sastre J.;  
 RT "Chicken serum albumin (Gal d 5\*) is a partially heat-labile inhalant  
 RT and food allergen implicated in the bird-egg syndrome";  
 RL Allergy 56:754-762(2001).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.  
 CC Partially heat-labile allergen that may cause both respiratory and  
 CC food-allergy symptoms in patients with the bird-egg syndrome.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; X60688; CAA43098.1; -;  
 DR EMBL; V00381; CAA23680.1; -;  
 DR PIR; S15571; ABCHS.  
 DR HSP; P02768; IE7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN\_3.  
 DR PROSITE; PS00012; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 23  
 FT CHAIN 24 615 SERUM ALBUMIN.  
 FT DOMAIN 24 209 ALBUMIN 1.  
 FT DOMAIN 216 401 ALBUMIN 2.  
 FT DOMAIN 408 599 ALBUMIN 3.  
 FT METAL 30 30 COPPER (BY SIMILARITY).  
 FT DISULFID 80 89 BY SIMILARITY.  
 FT DISULFID 102 118 BY SIMILARITY.  
 FT DISULFID 117 128 BY SIMILARITY.  
 FT DISULFID 152 197 BY SIMILARITY.  
 FT DISULFID 196 205 BY SIMILARITY.  
 FT DISULFID 228 274 BY SIMILARITY.  
 FT DISULFID 273 281 BY SIMILARITY.

FT DISULFID 293 307 BY SIMILARITY.  
 FT DISULFID 306 317 BY SIMILARITY.  
 FT DISULFID 344 389 BY SIMILARITY.  
 FT DISULFID 388 397 BY SIMILARITY.  
 FT DISULFID 420 466 BY SIMILARITY.  
 FT DISULFID 465 476 BY SIMILARITY.  
 FT DISULFID 489 505 BY SIMILARITY.  
 FT DISULFID 504 515 BY SIMILARITY.  
 FT DISULFID 542 587 BY SIMILARITY.  
 FT DISULFID 586 595 BY SIMILARITY.  
 FT CARBOHYD 500 500 N-LINKED (GLCNAC... (POTENTIAL).  
 FT CONFLICT 24 24 P -> M (IN REF. 3).  
 SQ SEQUENCE 615 AA; 69918 MW; 559E4BBAEC066C6 CRC64;  
 Query Match 50.2%; Score 1038.5; DB 1; Length 615;  
 Best Local Similarity 47.9%; Pred. No. 8.7e-70;  
 Matches 185; Conservative 75; Mismatches 125; Indels 1; Gaps 1;  
 QY 3 HKSEVAHREKDLGEENFKALVLAFAQYLQCCPFEDHVKLVNNEVTFEAKTCVADESAENC 62  
 DB 30 HKSEIAHYNDLKEETFKAVAMITFAQYLQRCSTEGLSKLVDVLDLAQKCVANEDAPEC 89  
 QY 63 DKSLHTLFGDKLCTVATLRETYGEMADCCAKQBFERNECFLOHKDDNPNLPR-LVRPEVD 121  
 DB 90 SKLPSIILDEIQVEKLRDSYGAMADCCSKADPERNECFLSFKVQSPDFVQVQRPASD 149  
 QY 122 VMCTAFHDNEETFLKKYLVEIARHPHYFYAPPELLFFAKRYKAFTTECCQADRAALLPK 181  
 DB 150 VICOEYQDNVSVFLGHFIYSVAERHPFLYAPAILSFVDFEHALQSCCKESDVGACLDTK 209  
 QY 182 LDELRLDEGKASSAKQLKLCASLOKFGERAPKAWAVARLSORFFKAPFAEYKLVTLTKV 241  
 DB 210 EIVMREKAGSVKQYVFCGLKQFGDRVFOARQLIYLQKYKAPFSEVSKFVHDSIGV 269  
 QY 242 HTECHGDLLECADDRADLANKYICENQDSISSKLKSCCKEPLLEKSHCAEVENDEMPAD 301  
 DB 270 HKSCCEGDVMYECNDDMARMMNSNLCSQDVFSGIKDCCKEPIVERSQCI MEAEFDEKPAD 329  
 QY 302 LPSLAADPFVSKDCKNYAEAKDVFGLFLEYARRHPDYSVLLRLAKTYETITLEKCC 361  
 DB 330 LPSLVEKYIIDKEVCKSFAGHDAPFAEFVYSRRHPFESQILINRIAKGYESLLEKCC 389  
 QY 362 AAADPHCYAKVDFEKPFLVEEPQL 387  
 DB 390 KTDNPAECYANACEQLNQHIKETQDV 415  
 RESULT 14  
 ALB2\_XENLA  
 ID ALB2\_XENLA STANDARD; PRT; 607 AA.  
 AC P14872;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 74 kDa serum albumin precursor.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 CX NCBI\_TaxID=8355;  
 [1]  
 RN SEQUENCE OF 3-607 FROM N.A.  
 RP MEDLINE=89313788; PubMed=2747653;  
 RA Moskaitis J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,  
 RA Schoenberg D.R.;  
 RT "Xenopus laevis serum albumin: sequence of the complementary  
 RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides  
 RT and the regulation of albumin gene expression by thyroid hormone  
 RL during development";  
 RN Mol. Endocrinol. 3:464-473(1989).  
 [2]  
 RP SEQUENCE OF 1-48 FROM N.A.  
 RX MEDLINE=88172470; PubMed=2451026;

RA Brownstein M.J., Usedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Iquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RN SEQUENCE OF 99-516 FROM N.A.  
RP MEDLINE=8216123; PubMed=2452956;  
RX Minghetti P.P., Law S.W., Dugaiczky A.,  
RA "The rate of molecular evolution of alpha-fetoprotein approaches that  
RT of pseudogenes";  
RL Mol. Biol. Evol. 2:347-358(1985).  
RN [5]  
RN SEQUENCE OF 477-551 FROM N.A.  
RP STRAIN=BALB/c;  
RX MEDLINE=90269606; PubMed=1971802;  
RA Boccaccio C., Deschattrette J., Meunier-Rotival M.,  
RT "Empty and occupied insertion site of the truncated LINE-1 repeat  
RT located in the mouse serum albumin-encoding gene";  
RL Gene 88:181-186(1990).  
RN [6]  
RN SEQUENCE OF 25-44.  
RP TISSUE=Liver;  
RX MEDLINE=93162044; PubMed=1286668;  
RA Giometti C.S., Taylor J., Tollaksen S.L.,  
RT "Mouse liver protein database: a catalog of proteins detected by two-  
RT dimensional gel electrophoresis";  
RL Electrophoresis 13:970-991(1992).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AJ011413; CAA09617.1; -  
DR EMBL; AK010025; BAB2650.1; -  
DR EMBL; BC049971; AAH49971.1; -  
DR EMBL; M16111; AAA37190.1; -  
DR EMBL; X13060; CAA31458.1; -  
DR PIR; A05339; A05139.  
DR HSPF; P02768; IEB7.  
DR SWISS-2DPAGE; P07724; MOUSE.  
DR MGD; MGI:87991; Alb1.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport prot.; 3.  
DR PRINTS; PR00802; SERUMALBUMIN  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SMC00103; ALBUMIN; 3.  
DR PROSITE; PS0012; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608 SERUM ALBUMIN.

| FT | DOMAIN   | 25      | 205       | ALBUMIN 1.              |
|----|----------|---------|-----------|-------------------------|
| FT | DOMAIN   | 212     | 397       | ALBUMIN 2.              |
| FT | DOMAIN   | 404     | 595       | ALBUMIN 3.              |
| FT | METAL    | 27      | 27        | COPPER.                 |
| FT | DISULFID | 77      | 86        | BY SIMILARITY.          |
| FT | DISULFID | 99      | 115       | BY SIMILARITY.          |
| FT | DISULFID | 114     | 125       | BY SIMILARITY.          |
| FT | DISULFID | 148     | 193       | BY SIMILARITY.          |
| FT | DISULFID | 192     | 201       | BY SIMILARITY.          |
| FT | DISULFID | 224     | 270       | BY SIMILARITY.          |
| FT | DISULFID | 269     | 277       | BY SIMILARITY.          |
| FT | DISULFID | 289     | 303       | BY SIMILARITY.          |
| FT | DISULFID | 302     | 313       | BY SIMILARITY.          |
| FT | DISULFID | 340     | 385       | BY SIMILARITY.          |
| FT | DISULFID | 384     | 393       | BY SIMILARITY.          |
| FT | DISULFID | 416     | 462       | BY SIMILARITY.          |
| FT | DISULFID | 461     | 472       | BY SIMILARITY.          |
| FT | DISULFID | 485     | 501       | BY SIMILARITY.          |
| FT | DISULFID | 500     | 511       | BY SIMILARITY.          |
| FT | DISULFID | 538     | 583       | BY SIMILARITY.          |
| FT | DISULFID | 582     | 591       | BY SIMILARITY.          |
| FT | CONFLICT | 27      | 27        | H -> D (IN REF. 6).     |
| FT | CONFLICT | 33      | 33        | H -> D (IN REF. 6).     |
| FT | CONFLICT | 41      | 41        | Q -> I (IN REF. 6).     |
| SQ | SEQUENCE | 608 AA; | 68692 MW; | 292F7C7EED3A61B4 CRC64; |

Query Match 76.6%; Score 1584; DB 1; Length 608;  
Best Local Similarity 71.8%; Pred. No. 2.8e-110; Indels 0; Gaps 0;  
Matches 278; Conservative 56; Mismatches 53;

| QY | 1   | DAHKEVAHRPKDGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE   | 60  |
|----|-----|--|-----|
| Db | 25  | BAHKSEIAHRYNDLGEQHFGLVLIATFQYLOKCSYDEHAKLVQEVTDFAKTCVADESAA  | 84  |
| QY | 61  | NCDKSLHLLFGDKLCTVATLRETYGEVADCCAKOPEPNECFLOHKDNNPLALVPEV     | 120 |
| Db | 85  | NCDKSLHLLFGDKLCAIPNLRENYGELADCTCKOPEPNECFLOHKDNNPLPPEPDEA    | 144 |
| QY | 121 | DMVCTAFHNDRETEFLKYLIEIARHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP | 180 |
| Db | 145 | EMACTSFENPTTFMGHYLHEVARHPYFYAPPELLFYAEQNEILLTCCAEADKESCLTP   | 204 |
| QY | 181 | KLDELRLDEGKASSAKORLKCSLOKFGEPKAPKAVARLSQTFPNADFAEITKLATDLTK  | 240 |
| Db | 205 | KLDGVKEALVSSVRQRMKSSMOKGERAPKAVARLSQTFPNADFAEITKLATDLTK      | 264 |
| QY | 241 | VTECHGDLLECCADRDADLAKYICENQDSISSKLECEKPELLEKSHCIAEVENDEMPA   | 300 |
| Db | 265 | VNKECHGDLLECCADRDADLAKYICENQDSISSKLECEKPELLEKSHCIAEVENDEMPA  | 324 |
| QY | 301 | DLPSLAADFVSKDKYKNAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLK      | 360 |
| Db | 325 | DLPSLAADFVSKDKYKNAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLK      | 384 |
| QY | 361 | CAAADPHECYAKVDFEPKPLVEEPNL                                   | 387 |
| Db | 385 | CAEAPPACVTVLAFQPLVEEPKQL                                     | 411 |

RESULT 13  
ALBU CHICK  
ID ALBU CHICK STANDARD; PRT; 615 AA.  
AC P19121;  
DT 01-NOV-1990 (Rel. 15, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).  
GN ALB.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_taxid=9031;

RN SEQUENCE FROM N.A.  
 RP STRAIN=MGS IDR; TISSUE=Liver;  
 RX MEDLINE=98116663; PubMed=9455485;  
 RA Yoshida K., Seto-Oshima A., Sinochara H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 RL DNA Res 4:351-354(1997).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AB006197; BAA21765.1; -;  
 DR PIR; JCS838; JCS838.  
 DR HSSP; P02768; 1E7B.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PD00802; SERUMALBUMIN.  
 DR PRODOM; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT PROPEP 19 24 BY SIMILARITY.  
 FT CHAIN 25 609 SERUM ALBUMIN.  
 FT DOMAIN 25 206 ALBUMIN 1.  
 FT DOMAIN 213 398 ALBUMIN 2.  
 FT DOMAIN 405 596 ALBUMIN 3.  
 FT METAL 28 28 COPPER.  
 FT DISULFID 78 87 BY SIMILARITY.  
 FT DISULFID 100 116 BY SIMILARITY.  
 FT DISULFID 115 126 BY SIMILARITY.  
 FT DISULFID 149 194 BY SIMILARITY.  
 FT DISULFID 193 202 BY SIMILARITY.  
 FT DISULFID 225 271 BY SIMILARITY.  
 FT DISULFID 270 278 BY SIMILARITY.  
 FT DISULFID 290 304 BY SIMILARITY.  
 FT DISULFID 303 314 BY SIMILARITY.  
 FT DISULFID 341 386 BY SIMILARITY.  
 FT DISULFID 385 394 BY SIMILARITY.  
 FT DISULFID 417 463 BY SIMILARITY.  
 FT DISULFID 462 473 BY SIMILARITY.  
 FT DISULFID 486 502 BY SIMILARITY.  
 FT DISULFID 501 512 BY SIMILARITY.  
 FT DISULFID 539 584 BY SIMILARITY.  
 FT DISULFID 583 592 BY SIMILARITY.  
 SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F67E7F1A48 CRC64;

Query Match 76.6%; Score 1585; DB 1; Length 609;  
 Best Local Similarity 73.6%; Pred. No. 2.4e-110;  
 Matches 284; Conservative 40; Mismatches 62; Indels 0; Gaps 0;

QY 2 AHKSEVAHFKDLGEENFALVLIAPAOYLQCPEDHVKLVNTEPAKTCVADESSEN 61  
 DB 27 AHKSEVAHFKDLGEENFALVLIAPAOYLQCPEDHVKLVNTEPAKTCVADESSEN 86  
 QY 62 COKSLHTLFGDKLCTVATIRYTGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEVD 121  
 DB 87 COKSLHTLFGDKLCTVATIRYTGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEVD 146

QY 122 VMCATFHDNBETFLKYLVEIARRHPYFYAPPELLFFANRYKAAFTCCQADKAAACLLPK 181  
 DB 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGFPELLYLDKTYAVLTTECCADDKACGLTPK 206  
 QY 182 LDEURDSGKASSAKORUKCASLOKFGERAFKANAVARLSORFPKAEFAEYKLVTDLTKV 241  
 DB 207 LDALKKALYSVAVRQRLKCSMKFGERAFKANAVARMSQTFPNADPAETKLTATDLTKV 266  
 QY 242 HTECHGDLLECADRADLAKYICENODSISKKLKECEKPLEKSHCIAEVENDEMPAD 301  
 DB 267 TQECHGDLLECADRADLAKYICENODSISKKLQACCKEMQLQSKSCLAEVEHDDMPAD 326  
 QY 302 LPSLAADFVSQCKVKNYAEAKVFLGMLYEVARRHPDYVSVLLLELAKTYETLEKCC 361  
 DB 327 LPALTADFVEDKQVKNYAEAKVFLGMLYEVARRHPDYVSVLLLELAKTYETLEKCC 386  
 QY 362 AAADPHECYAKVDFEPKPLVEEPONL 387  
 DB 387 AEADPHACYGHVDFEPKPLVEEPONL 412  
 RESULT 12  
 ALBU\_MOUSE  
 ID ALBU\_MOUSE STANDARD; PRT; 608 AA.  
 AC P07724; Q51802;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB OR ALBI OR ALB-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 EX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pease G., Quackenbush J.,  
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,



QY 1 DAHSEVAHFKDILGEENFKALVLIAPAOYLQCPEDHVKLVNVEYFAKTCVADSAE 60  
 DB 25 DTHKSEIAHFNLDGEENFQGLVLIAPSOYLQCPEDHVKLVNVEYFAKTCVADSAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATREYVYGEADCCAKQEPERNECFQHQDDNPRLPRVPEV 120  
 DB 85 GCDKSLHTLFGDELCKVATLREYVYGDVADCCQEPERNECFLNHKDDSPDLPKL-KPEP 143  
 QY 121 DVNCTAFHDEEFTLKKLYIEARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 144 DTLCAEPKADKKFVWGYLYEVARRHPYFYAPELLYANKYNGVFOCCQAEKGAACLLP 203  
 QY 181 KLDELREGKASAKORLKCAQKQGERAFKAWAVARLSORPKPAFAEVSKLVITDLTK 240  
 DB 204 KIDAMREKVLASSARQELRCASQKQGERALKAWVARLSQKPKADFTDVTIKVITDLTK 263  
 QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCECKPILKSHGICIAEVNDMPA 300  
 DB 264 VHKECCHGDLLECADRADLAKYICDHQDALSLSKLEKCCDKPYLEKSHGICIAEYDKDAVPE 323  
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360  
 DB 324 NLPLTADPADFEDKVCNQEAKDVFLGFLYSRRHPYAVSVLRLRLAKTYEATLEDC 383  
 QY 361 CAAADPEHYAKVDFEKPILVEFPQNL 387  
 DB 384 CAKEDPHACVATVDFKLVHVPQNL 410

## RESULT 9

ALBU RAT ID ALBU RAT STANDARD; PRT; 608 AA.  
 AC P02770; P11382;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].  
 GN ALB.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8123722; PubMed=7017712;  
 RA Sargent T.D., Yang M., Bonner J.;  
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).  
 RN [2]  
 RP SEQUENCE OF 1-38, AND PROCESSING.  
 RX MEDLINE=77249657; PubMed=893447;  
 RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-  
 piece. Analysis of the direct translation product of albumin  
 messenger RNA.";  
 RL J. Biol. Chem. 252:6846-6855(1977).  
 RN [3]  
 RP SEQUENCE OF 25-222.  
 RX MEDLINE=78109429; PubMed=564345;  
 RA Isemura S., Ikenaka T.;  
 RT "Amino acid sequences of fragments I and II obtained by cyanogen  
 bromide cleavage of rat serum albumin.";  
 RL J. Biochem. 83:35-48(1978).  
 RN [4]  
 RP SEQUENCE OF 223-288 AND 572-608.  
 RX MEDLINE=78260153; PubMed=956149;  
 RA Isemura S., Ikenaka T.;  
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
 the amino acid sequences of four fragments.";  
 RL J. Biochem. 79:1183-1196(1976).  
 RN [5]  
 RP SEQUENCE OF 166-174.  
 RC TISSUE=Plasma;

RX MEDLINE=87194805; PubMed=2437111;  
 RA Caraway R.E., Mitra S.P., Cochran D.E.;  
 RT "Structure of a biologically active neurotensin-related peptide  
 obtained from pepsin-treated albumin(s).";  
 RL J. Biol. Chem. 262:5968-5973(1987).  
 RN [6]  
 RP COPPER-BINDING.  
 RX MEDLINE=79001617; PubMed=80265;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
 RL Cancer Res. 38:3483-3486(1978).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and  
 blood flow (potential).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch)  
 CC -----  
 CC EMBL: V01222; CAA24532.1; -.  
 DR PIR: A93872; ABRYS.  
 DR HSP: P02768; IE7B.  
 DR InterPro: IPR00264; Serum albumin.  
 DR Pfam: PF00273; transport prot; 3.  
 DR PRINTS: PR00802; SERUMALBUMIN.  
 DR ProDom: PD002486; Serum albumin; 1.  
 DR SMART: SM00103; ALBUMIN\_3.  
 DR PROSITE: PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608  
 FT PEPTIDE 156 174  
 FT DOMAIN 25 205  
 FT DOMAIN 212 397  
 FT DOMAIN 404 595  
 FT METAL 27 27  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 148 193  
 FT DISULFID 192 201  
 FT DISULFID 224 270  
 FT DISULFID 269 277  
 FT DISULFID 289 303  
 FT DISULFID 302 313  
 FT DISULFID 340 385  
 FT DISULFID 384 393  
 FT DISULFID 416 462  
 FT DISULFID 461 472  
 FT DISULFID 485 501  
 FT DISULFID 500 511  
 FT DISULFID 538 583  
 FT DISULFID 582 591  
 FT VARIANT 282 262  
 FT CONFLICT 174 174  
 FT SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;  
 Y -> L (IN REF. 5).  
 Y 77.9%; Score 1610; DB 1; Length 608;  
 Best Local Similarity 72.1%; Pred. No. 3.4e-112;  
 Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;

Query Match  
 Best Local Similarity 72.1%; Pred. No. 3.4e-112;  
 Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;

```

Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
Polymorphism.
KW 1 18
KW SIGNAL 19 24
  FT PROPEP 25 607
  FT CHAIN 25 204
  FT DOMAIN 211 396
  FT DOMAIN 403 594
  FT METAL 27 27
  FT DISULFID 77 86
  FT DISULFID 99 115
  FT DISULFID 114 125
  FT DISULFID 147 192
  FT DISULFID 151 200
  FT DISULFID 223 269
  FT DISULFID 268 276
  FT DISULFID 288 302
  FT DISULFID 301 312
  FT DISULFID 339 384
  FT DISULFID 393 392
  FT DISULFID 435 461
  FT DISULFID 450 471
  FT DISULFID 484 500
  FT DISULFID 499 510
  FT DISULFID 537 582
  FT DISULFID 581 590
  FT VARIANT 214 214
  FT CONFLICT 302 302
  FT CONFLICT 304 305
  FT CONFLICT 324 324
  FT CONFLICT 394 395
  FT CONFLICT 437 437
  FT CONFLICT 493 494
  FT CONFLICT 494 494
  SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;

A -> T.
C -> K (IN REF. 6).
KP -> PC (IN REF. 6).
N -> D (IN REF. 6).
ST -> TS (IN REF. 6).
K -> R (IN REF. 12).
SE -> ES (IN REF. 6).

Query Match 78.5%; Score 1622.5; DB 1; Length 607;
Best Local Similarity 75.2%; Pred. No. 4e-113;
Matches 291; Conservative 46; Mismatches 49; Indels 1; Gaps 1

QY 1 DAHKSEVAHRPKDLGEENFKALVLIATAQYLQCCPFDPDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DTHKSEIAHRFKDLGEHFKGLVLIATFQYLQCCPFDPDHVKLVNEVTEFAKTCVADES 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 85 GCKSLHTLFGDELCKVASLRETYGDMADCCCKQEPERNECFLSHKDDSPDLPKL-KPDP 143
QY 121 DWMTAFHNDNEFTPLKKVLYETARRHPFYAPPELLFFAKYKAAFTCCOARDKAACLLP 180
DB 144 NTLCDSEFKADEKKFWGKLYETARRHPFYAPPELLYANKYNGVFOCCOARDKGACLLP 203
QY 181 KLDELDEGKASSAKQRLKCAASLOKQGERAFKAWAVARLSQRPPKAEFAEVSXLVDTLTK 240
DB 204 KIETWEEKVLASSARQRLRCAISQIKQGERALKANSVARLSQKTFPKAEFVEVTKLVDTLTK 263
QY 241 VHTCECHGDLSCADRDADLAKYICENODSTSSKLEKCECKPILLESKHCHIAEVENDEMPA 300
DB 264 VHKECHGDLSCADRDADLAKYICDNQDTTSSKLEKCCDKPILLESKHCHIAEYKDAIPE 323
QY 301 DLPLSLAADFVSEKDKVCKNYAEAKDVFGLQWFLYEYARRHPDYSVVLLIRLAKTYETTLK 360
DB 324 NUPPLTADPAEDKDKVCKNYQENKQAFGLSFLYEYSRRHPEYANSVLLRLAKEYEATLE 383
QY 361 CAAADPHECYAKVFDEPKPLVEEPONL 387
DB 384 CAKDDPHACYSTVFDKLKHLVDEPONL 410

RESULT 8
ID ALBU_SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)

```

Matches 291; Conservative 49; Mismatches 46; Indels 1; Gaps 1;  
 QY 1 DAKSEVAHRFKDGLGEENKALVLIAPQYLOQCFEDHVKLVNEVTFBFAKTCVADESAG 60  
 DB 23 DTVKSBIAHRFKDGLGSOYFKGLVLIAPSOHLQOCPEYEHVKLVREVTBFAKTCVADESAG 82  
 QY 61 NCDKSLHTLPGDKLCTVATILRETYGEMADCCAKQPERNECFLOHKNPNLPRIVREPV 120  
 DB 83 NCDKSLHTLPGDKLCAIPSLRHYGLADCCBPERNECFLOHKNPNLPRIVREPV 141  
 QY 121 DVNCTAFHNEETFLKKLYELIARRHPYFAPPELLAFKAKRYKAAFTCCQAAADKAACLLP 180  
 DB 142 VALCADFQDEQKFWGKYLIELIARRHPYFAPPELLAFKAKRYKAAFTCCQAAADKAACLLP 201  
 QY 181 KLDELDEGKASSAKORLKCASLOKGERAFAPAWAVARLSQPPKAEVSKLVTDLTK 240  
 DB 202 KIEHLREKVLTSAAKORLKCASLOKGERAFAPAWAVARLSQPPKAEVSKLVTDLTK 261  
 QY 241 VHTCECHGDLLECCADRLAKVICENQDISKKECKEPLLEKSHCIAEVENDEMPA 300  
 DB 262 VHTCECHGDLLECCADRLAKVICENQDISKKECKEPLLEKSHCIAEVENDEMPA 321  
 QY 301 DPLSLAADFVESKDVCKNTAENKADVFLGMLFLEYARRHPDYVLLLRLLAKTYETLEKC 360  
 DB 322 DLNPLEHDFVEDKEVCKNKEADVFLGMLFLEYARRHPDYVLLLRLLAKTYETLEKC 381  
 QY 361 CAAADPHCEYAKVDFEKFPLVBPQNL 387  
 DB 382 CAXEDFPACVATVDFKQFLVBPQNL 408  
 RESULT 7  
 ID ALBU BOVIN STANDARD; PRT; 607 AA.  
 AC P22789; Q02787;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1996 (Rel. 33, last sequence update)  
 DT 10-OCT-2003 (Rel. 42, last annotation update)  
 DE Serum albumin precursor (Allergen Bos d 6).  
 GN ALB.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Holowachuk E.W., Stoltenberg J.K., Reed R.G., Peters T. Jr.;  
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RP Tissue=Liver;  
 RC Barry T., Power S., Gannon F.;  
 RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP Tissue=Liver;  
 RC Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;  
 RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RA Wu H.T., Huang M.C.;  
 RP "the complete cDNA sequence of bovine serum albumin";  
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE OF 1-32.  
 RX MEDLINE=80024278; PubMed=488109;  
 RA McGillivray R.T.A., Chung D.W., Davie E.W.;  
 RT Biosynthesis of bovine plasma proteins in a cell-free system. Amino-  
 terminal sequence of prealbumin.  
 RP Eur. J. Biochem. 98:477-485(1979).  
 RN [6]  
 RN SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.  
 RP

RA Brown J.R.;  
 RT "Structure of bovine serum albumin";  
 RL Fed. Proc. 34:591-591 (1975).  
 RN [7]  
 RP REVISIONS TO 190-195.  
 RA Brown J.R.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [8]  
 RP SEQUENCE OF 402-433.  
 RX MEDLINE=82023364; PubMed=7283978;  
 RA Reed R.G., Putnam F.W., Peters T. Jr.;  
 RT "Sequence of residues 400-403 of bovine serum albumin";  
 RL Biochem. J. 191:867-868 (1980).  
 RN [9]  
 RP SEQUENCE OF 19-28.  
 RX MEDLINE=77134075; PubMed=843354;  
 RA Patterson J.B., Geller D.M.;  
 RT "Bovine microsomal albumin: amino terminal sequence of bovine  
 prealbumin";  
 RL Biochem. Biophys. Res. Commun. 74:1220-1226 (1977).  
 RN [10]  
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
 RX MEDLINE=91083649; PubMed=2260975;  
 RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;  
 RT "Rapid confirmation and revision of the primary structure of bovine  
 serum albumin by ESIMS and Frit-FAB LC/MS";  
 RL Biochem. Biophys. Res. Commun. 173:639-646 (1990).  
 RN [11]  
 RP SEQUENCE OF 25-41.  
 RX MEDLINE=88267456; PubMed=3389500;  
 RA Hsieh J.C., Lin F.P., Tam M.F.;  
 RT "Electroblotting onto glass-fiber filter from an analytical  
 isoelectrofocusing gel: a preparative method for isolating proteins  
 for N-terminal microsequencing";  
 RL Anal. Biochem. 170:1-8 (1988).  
 RN [12]  
 RP SEQUENCE OF 437-451.  
 RA Vilbois F.;  
 RL Submitted (AUG-1998) to Swiss-Prot.  
 RN [13]  
 RP DISULFIDE BONDS.  
 RA Brown J.R.;  
 RT "Structure of serum albumin: disulfide bridges";  
 RL Fed. Proc. 33:1389-1389 (1974).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M73993; AAA51411.1; -;  
 DR EMBL; X59889; CAA41735.1; -;  
 DR EMBL; Y17769; CAA76847.1; -;  
 DR EMBL; AF542068; MAN17824.1; -;  
 DR HSP; P02768; IE7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transprot\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.

```
DR PIR; S34053; ABHOS.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT SEQUENCE 607 AA; 68598 MW; 256P6E830A1B90C5 CRC64;

Query Match
Best Local Similarity 80.3%; Score 1660.5; DB 1; Length 607;
Matches 299; Conservative 40; Mismatches 47; Indels 1; Gaps 1;

Qy 1 DAHSEVAHFKD; GREENEALVLIAPAOYLQCCPREDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DTHKSEIHRNDIGERHFKGLVLFVAFSQYLQCCPFEDHVKLVNEVTEFAKCADESAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKDDNPNLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKDDNPNLPLVRPEV 143

Qy 121 DVMCTAFHDNEETFLKYLVEIARRHPEYFYAPPELLFYAKRYKAAFTCCCAADKACLIP 180
Db 144 DAQCAAFQEDPKFLGKLYLEVARRHPEYFYGPPELLFYAEYKADFTCCPADDKLACLIP 203

Qy 181 KLDELRLDGKASSAKQRLKCAASQKFGERAPKAWAVARLSQRFPAEVSXKLVTDLT 240
Db 204 KLDALKERILLSSAKERLKCSSFQNGERAVKAWAVARLSQKFPKADFAEVSKIIVTDLT 263

Qy 241 VHTCCCHGDLLECADRDADLAKYICENQDSISSKLECCCKPILLESKSHCIAVENDEMPA 300
Db 264 VHKECCHGDLLECADRDADLAKYICENQDSISSKLECCCKPILLESKSHCIAVENDEMPA 323

Qy 301 DLPSLAADFVSKCKYNAKADKDFVLMFLYIYARHPDYVSVLLLRLLAKTVETTLK 360
Db 324 DLPALAADFADKEICKHYKADKDFVLTGTFYIYARRHPDYVSVLLLRLLAKTVETTLK 383

Qy 361 CAADDPHCYAKVDFEFPLVEEPQNL 387
Db 384 CAADDPACVYRVDFQFPLVEEPKSL 410

RESULT 6
ALBU_PIG
ID ALBU_PIG
AC P08835; Q29019;
DT 01-NOV-1988 (Rel. 09, Created)

PIR; S34053; ABHOS.
HSP; P02768; IE7B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
SEQUENCE FROM N.A.
Tissue=Liver;
MEDLINE=89016582; PubMed=3174440;
Baldwin G.S. Weinstein J.;
"Nucleotide sequence of porcine liver albumin.";
Nucleic Acids Res. 16:9045-9045(1988).
FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
SUBCELLULAR LOCATION: Secreted.
TISSUE SPECIFICITY: Plasma.
SIMILARITY: Belongs to the ALB/APP/VDB family.
SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X12422; CAA30970.1; -.
EMBL; M36787; AAA30988.1; -.
PIR; S01392; ABEGS.
HSP; P02768; IE7B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
NON_TER 1 1
SIGNAL <1 16 BY SIMILARITY.
PROPEP 17 22 BY SIMILARITY.
CHAIN 23 605 SERUM ALBUMIN.
DOMAIN 23 202 ALBUMIN 1.
DOMAIN 209 394 ALBUMIN 2.
DOMAIN 401 592 ALBUMIN 3.
METAL 31 31 COPPER (BY SIMILARITY).
DISULFID 75 84 BY SIMILARITY.
DISULFID 97 113 BY SIMILARITY.
DISULFID 112 123 BY SIMILARITY.
DISULFID 145 190 BY SIMILARITY.
DISULFID 189 198 BY SIMILARITY.
DISULFID 221 267 BY SIMILARITY.
DISULFID 286 300 BY SIMILARITY.
DISULFID 299 310 BY SIMILARITY.
DISULFID 337 382 BY SIMILARITY.
DISULFID 381 390 BY SIMILARITY.
DISULFID 413 459 BY SIMILARITY.
DISULFID 458 469 BY SIMILARITY.
DISULFID 482 498 BY SIMILARITY.
DISULFID 497 508 BY SIMILARITY.
DISULFID 535 580 BY SIMILARITY.
DISULFID 579 588 BY SIMILARITY.
CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1P4FF CRC64;

Query Match
Best Local Similarity 78.5%; Score 1623.5; DB 1; Length 605;
Pred. No. 3.3e-113;
Pred. No. 3.3e-113;
```





```

ID AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OC NCBI_TaxID=9685;
RN [1]
RX MEDLINE=96194824; PubMed=8647469;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSP; P02768; IEB7.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PD00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 383 BY SIMILARITY.
FT DISULFID 384 395 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F6085F CRC64;
Query Match 83.5%; Score 1727; DB 1; Length 608;
Best Local Similarity 80.6%; Pred. No. 7e-121;

```

```

Matches 312; Conservative 34; Mismatches 41; Indels 0; Gaps 0;
QY 1 DAHSEVAHRRFKDLGRENFKALVLIAPAOVLQOCPPEDHVKLVNEVTEPAKTCVADSEAE 60
DB 25 EAHQSEIAHRRFNDLGEHFRGLVLAFAFSQYLQOCPPEDHVKLVNEVTEPAKTCVADSEAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 85 NCEKSLHLLGDKLCTVASLRDKYGEWADCCCKEKEPERNECFLOHKDDNPNLPRLVPEV 144
QY 121 DVMTAFHDNBEETPLKYLIEIARRHPYFYAPPELLFPKRYKAAFTCCCAADKAAACLLP 180
DB 145 DAMCTAFHENEQREFGLKYLIEIARRHPYFYAPPELLFPKRYKAAFTCCCAADKAAACLLP 204
QY 181 KLDELREGKASSAKQRLKASLOKFGERAFKAWARLSORPPKAFPAFVSLVLTDLTK 240
DB 205 KVDALREKYLASSAKERLKASLOKFGERAFKAWARLSORPPKAFPAFVSLVLTDLTK 264
QY 241 VHTCCGHDLLLEADDRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
DB 265 IHKCCGHDLLLEADDRADLAKYICENQDSISTKLKCCGKPVLEKSHCHISEVERDELPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEVARRHPDYVVLLLRLLAKTVETTLK 360
DB 325 DLPLAVDFVEDKVKYQKADVFLGTFLYEVSRHRHPYSVLLRLAKEVZATLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQL 387
DB 385 CATDDPPACYAHVDFEFPKPLVEEPHNL 411
RESULT 4
ALBU CANFA STANDARD; PRT; 608 AA.
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandejsky-Fichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504912;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).

```

```

Query Match      100.0%; Score 2068; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 3.5e-146;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHFRFKDLGSENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHFRFKDLGSENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRIVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRIVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDMPA 300
DB 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDMPA 324
QY 301 DLPSLAADPVESKDVCKNVAEKDVLGMLFYEARHHPDYSVVLLRLAKTYETTLTK 360
DB 325 DLPSLAADPVESKDVCKNVAEKDVLGMLFYEARHHPDYSVVLLRLAKTYETTLTK 384
QY 361 CAADPHECYAKVDFBPKPLVEEPQNL 387
DB 385 CAADPHECYAKVDFBPKPLVEEPQNL 411

RESULT 2
ALBU_MACMU STANDARD; PRT; 600 AA.
ID ALBU_MACMU
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
RT in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; M90463; AAA36906.1; -.
DR F01; A47391; A47391.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot. 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1
FT SIGNAL <1 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match      96.1%; Score 1988; DB 1; Length 600;
Best Local Similarity 94.6%; Pred. No. 2.9e-140;
Matches 366; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DAHKSEVAHFRFKDLGSENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 17 DTHKSEVAHFRFKDLGSENFKGLVLAFAVSYLQCCPEEHEVHKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRIVRPEV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRIVRPEV 136
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 137 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAARYKAAFAECCQAADKAAACLLP 196
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240
DB 197 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 256
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDMPA 300
DB 257 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDMPA 316
QY 301 DLPSLAADPVESKDVCKNVAEKDVLGMLFYEARHHPDYSVVLLRLAKTYETTLTK 360
DB 317 DLPSLAADPVESKDVCKNVAEKDVLGMLFYEARHHPDYSVVLLRLAKTYETTLTK 376
QY 361 CAADPHECYAKVDFBPKPLVEEPQNL 387
DB 377 CAADPHECYAKVDFBPKPLVEEPQNL 403

RESULT 3
ALBU_FELCA

```

RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Hilton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RN SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RN SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.O.;  
RL (In) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RN SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RN SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RN SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RN SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RN DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RN BIIRUBIN-BINDING SITE.  
RA Jacobsen C.;  
RX MEDLINE=78186630; PubMed=656055;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).

RN [16]  
RN VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RN VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satch C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioke N., Satch C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese alloalbumins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RN VARIANTS MANAUS; OSAKA; NAGOVA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RN DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
human serum albumin whose precursor has an aberrant signal peptidase  
cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RN VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y., Anaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
RN [24]  
RN VARIANT CASEBOOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
albumin: albumin Casebrook (494 Asp-->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two alloalbumins with identical electrophoretic mobility are produced

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 28.2299 Seconds  
(without alignments)  
713.823 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_1\_387  
Perfect score: 2068  
Sequence: 1 DAHSEVAHRFKDLEENFX.....ECYAKVDFEKFPLVEEPQNL 387

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1          | 2068   | 100.0       | 609    | 1  | ALBU HUMAN  |
| 2          | 1988   | 96.1        | 600    | 1  | ALBU MACMU  |
| 3          | 1727   | 83.5        | 608    | 1  | ALBU FELCA  |
| 4          | 1690   | 81.7        | 608    | 1  | ALBU CANFA  |
| 5          | 1660.5 | 80.3        | 607    | 1  | ALBU HORSE  |
| 6          | 1623.5 | 78.5        | 605    | 1  | ALBU PIG    |
| 7          | 1622.5 | 78.5        | 607    | 1  | ALBU BOVIN  |
| 8          | 1613.5 | 78.0        | 607    | 1  | ALBU SHEEP  |
| 9          | 1610   | 77.9        | 608    | 1  | ALBU RAT    |
| 10         | 1585   | 76.6        | 608    | 1  | ALBU RABIT  |
| 11         | 1585   | 76.6        | 609    | 1  | ALBU MERUN  |
| 12         | 1584   | 76.6        | 608    | 1  | ALBU MOUSE  |
| 13         | 1038.5 | 50.2        | 615    | 1  | ALBU CHICK  |
| 14         | 766    | 37.0        | 607    | 1  | ALB2 XENLA  |
| 15         | 726.5  | 35.1        | 606    | 1  | ALB1 XENLA  |
| 16         | 724.5  | 35.0        | 609    | 1  | FETA PANTR  |
| 17         | 721.5  | 34.9        | 609    | 1  | FETA GORGO  |
| 18         | 720.5  | 34.8        | 609    | 1  | FETA HUMAN  |
| 19         | 691    | 33.4        | 599    | 1  | AFAM HUMAN  |
| 20         | 675    | 32.6        | 609    | 1  | FETA HORSE  |
| 21         | 671    | 32.4        | 611    | 1  | AFAM MOUSE  |
| 22         | 656    | 31.7        | 608    | 1  | AFAM RAT    |
| 23         | 615    | 29.8        | 605    | 1  | FETA MOUSE  |
| 24         | 603    | 29.2        | 611    | 1  | FETA RAT    |
| 25         | 468    | 22.6        | 608    | 1  | ALB1_SALSA  |
| 26         | 465    | 22.5        | 608    | 1  | ALB2_SALSA  |
| 27         | 357    | 17.3        | 474    | 1  | VTDB HUMAN  |
| 28         | 353    | 17.1        | 382    | 1  | ALBU RANCA  |
| 29         | 348    | 16.8        | 476    | 1  | VTDB RAT    |
| 30         | 345    | 16.7        | 476    | 1  | VTDB RABIT  |
| 31         | 341    | 16.5        | 476    | 1  | VTDB MOUSE  |
| 32         | 311    | 15.0        | 1423   | 1  | ALBU_PETWA  |
| 33         | 133    | 6.4         | 1605   | 1  | RRB1_MOUSE  |

34 122.5 5.9 344 1 MST1\_DROHY  
35 111 5.4 40 1 ALB1\_TRASC  
36 108 5.2 1410 1 RRB1\_HUMAN  
37 107.5 5.1 1228 1 YHV5\_YEAST  
38 106 5.1 650 1 BM86\_BOOMI  
39 102 4.9 1391 1 MST2\_DROHY  
40 101.5 4.9 3038 1 TRIO\_HUMAN  
41 100 4.8 275 1 MST3\_DROHY  
42 99.5 4.8 1962 1 MYSA\_DROME  
43 99 4.8 1285 1 SLI7\_ENTHI  
44 99 4.8 2867 1 RRB2\_PLAUV  
45 99 4.8 8797 1 SNE1\_HUMAN

Q08695 drosophila  
P81188 trachemys s  
Q922e9 homo sapien  
P38851 saccharomyc  
F20736 boophilus m  
Q08696 drosophila  
O75962 homo sapien  
O01395 drosophila  
P05661 drosophila  
P23502 entamoeba h  
Q00799 plasmodium  
Q8n191 homo sapien

#### ALIGNMENTS

RESULT 1  
ALBU HUMAN  
ID ALBU\_HUMAN STANDARD; PRT; 609 AA.  
AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9JUZO;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RA Beattie W.G., Dugaiczky A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=82081882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczky A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RX Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.

us-09-832-929-18\_copy\_1\_387.rpr

Mon Apr 19 13:27:13 2004

Search completed: April 19, 2004, 12:02:20  
Job time : 48.169 secs

Db 266 KDCCHDDMFECWTERLELSEHTCQHKDELSTKLEKCCNPLDERTYCIIVLENDVPAEL 325  
Qy 303 PSIAADPVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYVSVLLARLAKTYETTTLEKCA 362  
Db 326 SQPITEFTEDPHVCYAEKNEVFLGRYLHVAVRKEHQELSEQFLLOSAKEYESLLNKCK 385  
Qy 363 AADPHCEYAKVDFEF 377  
Db 386 TDNPPECYKDGADRF 400

RESULT 14  
ABXL68  
68K serum albumin precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 28-Apr-1995  
C:Accession: A41682; S02692  
R:Meskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.  
Mol. Endocrinol. 3, 464-473, 1989  
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.  
A:Reference number: A41682; MUID:99313788; PMID:2747653  
A:Accession: A41682  
A:Molecule type: mRNA  
A:Residues: 1-608 <MOS>  
A:Cross-references: GB:M18350  
R:Schorpp, M.; Doebbeling, U.; Wagner, U.; Ryffel, G.U.  
J. Mol. Biol. 199, 83-93, 1988  
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deleted  
A:Reference number: S02692; MUID:98172470; PMID:2451026  
A:Accession: S02692  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-48 <SCH>  
A:Cross-references: EMBL:Z26825  
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyz  
C:Genetics:  
A:Introns: 27/1  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:23-608/Domain: 68K serum albumin #status predicted <MAT>  
F:32-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>  
F:30/Binding site: copper (His) #status predicted  
F:80-89,102-118,117-128,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,  
Query Match 35.78; Score 737.5; DB 1; Length 608;  
Best Local Similarity 36.9%; Pred. No. 2.5e-47;  
Matches 139; Conservative 72; Mismatches 159; Indels 7; Gaps 2;

Qy 3 HKSEVAHRFDLGBENFKALVLIAPAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAENC 62  
Db 30 HHKHIADWYLLTERTFKGLTAVISQNLQKSLSEELSKLVNEINDFAKCTGNDKTPTEC 89  
Qy 63 DKSLLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHK--DDNPRLVLRPEV 120  
Db 90 EKPITGLFYDKLADPKVGVNYESKESCKQDPERACQCFRAHRVFEHN-----VRPKP 144  
Qy 121 DVMCTAFHDNEETPLKKYLEIARRHPYFVAPPELLFFAKRYKAATFECQADKAAACLLP 180  
Db 145 EETCALPKEHPDILLSAFIHEARNHDPYPPAVLLITQYGLVHCCEEDKDKCFAE 204  
Qy 181 KLDELDEGKASSAKQRLKCAQKQFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240  
Db 205 KMKELMKHSHSIEDKQHFQWVNNYPERVIRKALNARVSHRYPKDPFKLAHFTTBETH 264  
Qy 241 VHTCCCHDLEECADRADLAKYICENODSTSSKLKCECKEPKLEKSHCHIAEVENDEMPA 300

Db 265 FIKDCHGDMFECWTERLELSEHTCQHKDELSTKLEKCCNPLDERTYCIIVLENDVPA 324  
Qy 301 DLPSLAADPVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYVSVLLARLAKTYETTTLEK 360  
Db 325 ELSKPITEFTEDPHVCYAEKNEVFLGRYLHVAVRKEHQELSEQFLLOSAKEYESLLNK 384  
Qy 361 CAAADPHCEYAKVDFEF 377  
Db 385 CFSDNFPECYKDGADRF 401

RESULT 15  
JC4258  
alpha-fetoprotein precursor - chimpanzee  
C:Species: Pan troglodytes (Chimpanzee)  
C:Date: 27-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 20-Aug-1999  
C:Accession: JC4258  
R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.  
Gene 162, 213-220, 1995  
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to  
A:Reference number: JC4258; MUID:96032345; PMID:7557431  
A:Accession: JC4258  
A:Molecule type: DNA  
A:Residues: 1-609 <NIS>  
A:Cross-references: GB:U21916; NID:9841311; PIDN:AAA91641.1; PID:9841312  
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver ar  
o similar properties and structure.  
C:Genetics:  
A:Gene: afp  
A:Map position: 3p  
A:Introns: 29/1, 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 5  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: glycoprotein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-609/Domain: alpha-fetoprotein #status predicted <MAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>  
F:42,251/Binding site: carboxylate (Asn) (covalent) #status predicted  
Query Match 35.0%; Score 724.5; DB 2; Length 609;  
Best Local Similarity 36.3%; Pred. No. 2.3e-46;  
Matches 142; Conservative 75; Mismatches 167; Indels 7; Gaps 3;

Qy 3 HKSE-----VAHRFDLGBENFKALVLIAPAQYLOQCFFEDHVKLVNEVTEFAKTCVADE 57  
Db 22 HRNEYGIASILDYSQCTAEINLTDLATIFFAQFQERATYKEVSKVWDALTAIEKPTGDE 81  
Qy 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNP-NLRLV 116  
Db 82 QSAGCLENLPAFLLEELCREKEILEKYGH-SDCCSQSEEGRHNCFLAHKKPTASIPFPQ 140  
Qy 117 RPEVDMCTAFHDNEETPLKKYLEIARRHPYFVAPPELLFFAKRYKAATFECQADKAA 176  
Db 141 VPEPVTSCAEYEEERETPMNKFIVEIARRHPFLFAPTILLWAARYDKLIIPCCKAEAVE 200  
Qy 177 CLLPKLDELDEGKASSAKQRLKCAQKQFGERAFKAWAVARLSQRPFAEFAEVSKLVT 236  
Db 201 CFQTKAATVTKELRESSLLNHACAVMKNFGTRIFQAITVTKLSQKTKVNFTEIQKLV 260  
Qy 237 DLTKVHTCCCHDLEECADRADLAKYICENODSTSSKLKCECKEPKLEKSHCHIAEVEN 296  
Db 261 DVAVHVEHCCRGDVLDCLDQGEKIMSYTCQQDSTLSNKITECCCKLTTLERQQCIHAEN 320  
Qy 297 EMPADLPISLAADFVBSKDVCKNYAEAKDVFGLGMFLYEVARRHPDYVSVLLLRLLAKTYETT 356  
Db 321 EKPEGLSPNLRFLGDRDNFQSSGKKNIFLASFVHEYSRRHPQLAVSVILRVAKGYQEL 380  
Qy 357 LEKCAAAADPHCEYAKVDFEFKPLVPEPQNL 387  
Db 381 LEKCFQTEPNLEPCQDKGEELQKVIQESQAL 411





A:Cross-references: EMBL:X13060; NID:g52939; PIDN:CAA1458.1; PID:G899334  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>  
F:123-296/Domain: serum albumin repeat homology <SA2>  
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match. 61.1%; Score 1264; DB 2; Length 453;  
Best Local Similarity 70.6%; Pred. No. 1.1e-86;  
Matches 221; Conservative 45; Mismatches 47; Indels 0; Gaps 0;

QY 75 CTVATLRETYGEMADCCAKPEPERNECFLOHQDNDPNLRLVRPEVDVWCTAFHDNERTF 134  
DB 1 CAPNLRENTGELADCTKPEPERNECFLOHQDNDPNLPPFERPEAEAMCISFKNPTTF 60  
QY 135 LKKYLVEIARRHPYFYAPELLFFPKRYKAAFTCCQAADKAACLLPKLDELDEGKASSA 194  
DB 61 MGHYLVHEVARHPYFYAPELLFYAEQNEILTQCABADKESCLTPKLDGVKEXALVSSV 120  
QY 195 KORLKCASLOKFGERAFAKAVAVELSRQFPKAEFAEVSKLVDTLTKVHTECHGDLLECA 254  
DB 121 RQRMKCSMQKFGERAFAKAVAVELSGTFPNADFAITKLATDLTKVKECHGDLLECA 180  
QY 255 DDRADLAKYICENODTSISKLKECCERPLLEKSHCIAEVDENDEMPADLPSLAADFVESKD 314  
DB 181 DDRAELAKYMCENQATISSKLQTCDDKPLLKKAHCLSEVEHDTWPAADLPAIAADFVEDQE 240  
QY 315 VCKNYAAKDVFLGNFLFYEVARRHPDYSVULLLRLAKYETTLKCCAAADPHECYAKVF 374  
DB 241 VCKNYAAKDVFLGTFLYEYSRRHPDYSVULLLRLAKYETTLKCCAAANPPACYGTVL 300  
QY 375 DEPKPLVEEPQNL 387  
DB 301 AEFQPLVEEPKQL 313

RESULT 11  
ABCS  
serum albumin precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S15571; A05078; A13451  
R:Caesady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.  
submitted to the EMBL Data Library, July 1991  
A:Reference number: S15571  
A:Accession: S15571  
A:Molecule type: mRNA  
A:Residues: 1-615 <CAS>  
A:Cross-references: EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748  
R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.  
J. Biol. Chem. 259, 4556-4564, 1983  
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein  
A:Reference number: A05078; MUID:83161037; PMID:6187737  
A:Accession: A05078  
A:Molecule type: DNA  
A:Residues: 1-28 <AC>  
A:Cross-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039  
R:Rosen, A.M.; Geller, D.M.  
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977  
A:Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.  
A:Reference number: A13451; MUID:78019943; PMID:911327  
A:Accession: A13451  
A:Molecule type: protein  
A:Residues: 19-23, M, 23-30 <ROS>  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, zinc, and iron. It is a carrier protein for many hormones and drugs. It promotes the transfer of many drugs across the membranes of the liver.  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-26/Domain: propeptide #status predicted <PRO>  
F:21-61/Product: serum albumin #status predicted <MAT>  
F:32-206/Domain: serum albumin repeat homology <SA1>  
F:225-398/Domain: serum albumin repeat homology <SA2>

F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-203/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4  
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 78.0%; Score 1613.5; DB 1; Length 607;  
Best Local Similarity 74.2%; Pred. No. 1.3e-112; Indels 1; Gaps 1;  
Matches 287; Conservative 48; Mismatches 51;  
QY 1 DAHSEVAHREKDLGEENFKALVLIAPAFYQYQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DTHKSEIARRNDLGEENFQGLVLIAPSYQLQCCPFDEHVKLVKELTEFAKTCVADESAA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLPLVRPEV 120  
DB 85 GCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLPLVRPEV 143  
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 144 DTLCAEFKADKKWGKLYIEVARRHPYFYAPPELLYANKYNGVFQSCQAEDKGAACLLP 203  
QY 181 KLDELDEGKASSAKQRLKCAKSLQKQGERAFKAWARLQRPFKAEFAEYVKLVTLDTLK 240  
DB 204 KIDAMREKVLASSARQRLRCASIQKFGERALKAWARLQKQFPKADFTDVTIVTLDTLK 263  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCHIAEVENDEMPA 300  
DB 264 VKECCCHGDLLECADRADLAKYICDHODALSSKKECCQKPLEKSHCHIAEVDKDAVPE 323  
QY 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYIEYARRHPDYVSVLLRLAKTYETTLK 360  
DB 324 NLPLTADFAEDKECKNYQAKDVFGLSGFLYIEYARRHPDYVSVLLRLAKTYETTLK 383  
QY 361 CAADPHCYAKVDFDEKPLVEEPQNL 387  
DB 384 KAKEDPHACYATVFDKHLVDEPQNL 410

## RESULT 8

ABRTS  
serum albumin precursor - rat  
N:Alternate names: preproalbumin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999  
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233  
R:Sargent, T.D.; Yang, M.; Bonner, J.  
Proc. Natl. Acad. Sci. U.S.A. 76, 243-246, 1981  
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
A:Reference number: A93872; MUID:81223722; PMID:7017712  
A:Accession: A93872  
A:Molecule type: mRNA  
A:Residues: 1-608 <SAR>  
A:Cross-references: GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628  
R:Staass, A.W.; Bennett, C.D.; Donchue, A.M.; Rodkey, J.A.; Alberts, A.W.  
J. Biol. Chem. 252, 6846-6855, 1977  
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis  
A:Reference number: A92211; MUID:77249657; PMID:893447  
A>Note: cleavages during protein maturation  
A:Accession: A92211  
A:Molecule type: protein  
A:Residues: 1-38 <STR>  
R:Isemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleaved  
A:Reference number: A91946; MUID:76109429; PMID:564345  
A:Accession: A91946  
A:Molecule type: protein  
A:Residues: 25-222 <IS1>  
R:Isemura, S.; Ikenaka, T.

J. Biochem. 79, 1183-1196, 1976  
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino  
A:Reference number: A91940; MUID:76260153; PMID:956149  
A:Accession: A91940  
A:Molecule type: protein  
A:Residues: 223-288;572-608 <IS2>  
A>Note: 262-Leu was also found  
R:Aoyagi, Y.; Ikenaka, T.; Ichida, P.  
Cancer Res. 38, 3483-3486, 1978  
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A:Reference number: A90738; MUID:79001617; PMID:80265  
A:Contents: annotation; copper binding  
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: C45800  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <CAR>  
R:Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987  
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improve  
A:Reference number: I57621; MUID:87286876; PMID:3475566  
A:Accession: I57621  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-5 <RES>  
A:Cross-references: GB:M16825; NID:9202828; PIDN:AAA40712.1; PID:9554412  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:18/Domain: signal sequence #status experimental <PRO>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-608/Product: serum albumin #status experimental <MAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status experimental  
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,

Query Match 77.9%; Score 1610; DB 1; Length 608;  
Best Local Similarity 72.1%; Pred. No. 2.4e-112;

Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;  
QY 1 DAHSEVAHREKDLGEENFKALVLIAPAFYQYQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHREKDLGEENFKALVLIAPAFYQYQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLPLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 EAMCTFQENPTSLFGLHVEARRHPYFYAPPELLFYAEKNEVLTQCTESDKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKCAKSLQKQGERAFKAWARLQRPFKAEFAEYVKLVTLDTLK 240  
DB 205 KLDVKEKALVAARQMKCSQVQGERAFKAWARLQRPFKAEFAEYVKLVTLDTLK 264  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCHIAEVENDEMPA 300  
DB 265 INKECCCHGDLLECADRADLAKYICENQATISSKLAQCCDPVLOKSOCLAEETHDNLPA 324  
QY 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYIEYARRHPDYVSVLLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYIEYARRHPDYVSVLLRLAKTYETTLK 384  
QY 361 CAADPHCYAKVDFDEKPLVEEPQNL 387  
DB 385 CAEGDPPACYGTVLAEFQPLVEEPQNL 411

C;Accession: A38885; A36401, A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94  
E;Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.  
submitted to the EMBL Data Library, August 1991  
A;Description: Bovine serum albumin: cDNA sequence and expression.  
A;Reference number: A38885  
A;Accession: A38885  
A;Molecule type: mRNA  
A;Residues: 1-607 <HOL>  
A;Cross-references: EMBL:M73215  
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
A;Reference number: A36401; MUID:91083649; PMID:2260975  
A;Accession: A36401  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>  
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979  
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A;Reference number: A91258; MUID:80024278; PMID:488109  
A;Accession: A91258  
A;Molecule type: protein  
A;Residues: 1-32 <MAG>  
R;Haieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A;Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing gel  
A;Reference number: A60808; MUID:88267456; PMID:3389500  
A;Accession: B60808  
A;Molecule type: protein  
A;Residues: 25-41 <HSI>  
R;Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A;Title: Tooth 'enamelin', identified mainly as serum proteins. Major 'enamelin' is albumin  
A;Reference number: S10780; MUID:90336641; PMID:2379503  
A;Accession: S10780  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>  
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1580-1584, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: D45800  
A;Molecule type: protein  
A;Residues: 163-172 <CAR>  
R;Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe  
A;Reference number: A26693; MUID:87194805; PMID:2437111  
A;Accession: A26693  
A;Molecule type: protein  
A;Residues: 165-172, 'L', <CA2>  
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A;Title: Sequence of residues 400-403 of bovine serum albumin.  
A;Reference number: A90309; MUID:82023364; PMID:7283978  
A;Accession: A90309  
A;Molecule type: protein  
A;Residues: 402-433 <REE>  
R;Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A;Title: Structure of bovine serum albumin.  
A;Reference number: A91458  
A;Accession: A91458  
A;Molecule type: protein  
A;Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'E'  
R;Brown, J.R.  
submitted to the Atlas, April 1975  
A;Reference number: A94551  
A;Accession: A94551  
A;Molecule type: protein  
A;Residues: 190-195 <BR2>  
R;Brown, J.R.  
Fed. Proc. 33, 1389, 1974

A;Reference number: A91457  
A;Contents: annotation; disulfide bonds  
R;Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
A;Reference number: S55232; MUID:95031935; PMID:7945219  
A;Accession: S55232  
A;Status: Preliminary  
A;Molecule type: protein  
A;Residues: 529-536; 569-572 <WER>  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; copper binding; duplication; plasma  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-607/Product: serum albumin #status experimental <MPT>  
F;29-201/Domain: serum albumin repeat homology <SA1>  
F;220-393/Domain: serum albumin repeat homology <SA2>  
F;412-591/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,  
Query Match 78.3%; Score 1618.5; DB 1; Length 607;  
Best Local Similarity 74.9%; Pred No. 5.6e-113;  
Matches 290; Conservative 46; Mismatches 50; Indels 1; Gaps 1;  
QY 1 DAHSEVAHRRFKDLGEEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DTHKSEIAHRRFKDLGEEQFKGLVLIATFQYLLQCCPFEDHVKLVNELTEFAKTCVADESHA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPENECEFLQHKDNDNPLRLVRPEV 120  
Db 85 GCKSLHTLFGDECKVASLRETYGDMADCCCKOEPENECEFLSHKDDSPOLPKL-KDDP 143  
QY 121 DVMCTAFHNDDETEFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180  
Db 144 NTLCDERKADKKFKGKYLIEIARRHPYFYAPPELLFYANKYGVFQDCQAEKDGACLLP 203  
QY 161 KLDELDEGKASSAKQRLKASLOKFGERAKAVANAVARLSORFPKABFAEYVKLVTDLT 240  
Db 204 KIETVREKVLASSARQURUCASIOKFGERALKAVARLSOKFPKAEFVEVTKLVTDLT 263  
QY 241 VHTCECHGDLLECDRADLAKYICENQDISSSKLKECEKFLLEKSHCIAEVENDEMPA 300  
Db 264 VKECCHGDLLECDRADLAKYICDNQDTISSKLKECCDKFLLEKSHCIAEVEKDAIPE 323  
QY 301 DLPSLAADFVSKDVKQYAAKQVFLGMFLYEVARHPDYVSVLLRLAKTYETTLK 360  
Db 324 NLPLTADFAEDKQYCKNYQEAQDAFLGSLFYYSRRHPYAVSVLLRLAKYEATLEEC 383  
QY 361 CAADPHCEYAKVDPDERKPLVEEPQNL 387  
Db 384 CAKDHPHACYTVDFDKLHLVDFQNL 410  
RESULT 7  
ABSHS  
serum albumin precursor - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: S06936  
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A;Reference number: S06936; MUID:90098888; PMID:2602160  
A;Accession: S06936  
A;Molecule type: mRNA  
A;Residues: 1-607 <SRO>  
A;Cross-references: EMBL:X17055; NID:gl386; PIDN:CAA34903.1; PID:gl387  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the mem  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; Plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>

QY 301 DLPSLAADPVESKDVCKNVAEAKDVLGMLFYIYARRHPDYSVLLRLRAKYETTTLEKC 360  
Db 325 DLPLAVDFVEDKVCNKYQKADKDFLGTFLYESSRRHPESVSLRLRAKYETTTLEKC 384  
QY 361 CAADPHCEYAKVDFDEKPLVBPQNL 387  
Db 385 CATDDPPACVAFVDFDEKPLVBPQNL 411

RESULT 4  
ABOS  
serum albumin precursor - horse  
C:Species: Equus caballus (domestic horse)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S34053  
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
F:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A:Reference number: S34053; MUID:93345495; PMID:8344282  
A:Accession: S34053  
A:Molecule type: mRNA  
A:Residues: 1-607 <HOA>  
A:Cross-references: GB:X74045; NID:G999671; PIDN:CAA52194.1; PID:G999672  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4  
F:263/Binding site: bilirubin (lys) #status predicted

Query Match 80.3%; Score 1660.5; DB 1; Length 607;  
Best Local Similarity 77.3%; Pred. No. 4.2e-116;  
Matches 299; Conservative 40; Mismatches 47; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
Db 25 DTHKSEIAHRFNDJGEKFGKGLVAFSQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLPLVRPEV 143  
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAAPTECCQADKAACLLP 180  
Db 144 DAQCAAFQEDPKFLGKLYIEVARRHPYFYGPPELLFAHEYKADFTCCPADDKLACLIP 203  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEYSKLVTDLT 240  
Db 204 KLDALKEILLSSAKERLKSSQNFGRVAVKANSVARLSQKFPKADFAEVSKIIVTDLTK 263  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 264 VHEKCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 323  
QY 301 DLPSLAADPVESKDVCKNVAEAKDVLGMLFYIYARRHPDYSVLLRLRAKYETTTLEKC 360  
Db 324 DLPALAAADFAEDKEICKHYKAKADVFLGTFLLYESSRRHPDYSVLLRLRIAKTYEATLEKC 383

QY 361 CAADPHCEYAKVDFDEKPLVBPQNL 387  
Db 384 CAADPPACVAFVDFQFTPLVEEPKSL 410

RESULT 5  
ABOS  
serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S01382; A61006  
R:Weinstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 9045, 1988  
A:Title: Nucleotide sequence of porcine liver albumin.  
A:Reference number: S01382; MUID:89016582; PMID:3174440  
A:Accession: S01382  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-605 <WEI>  
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798  
R:Lineback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral  
A:Reference number: A61006; MUID:89263769; PMID:2728927  
A:Accession: A61006  
A:Molecule type: protein  
A:Residues: 23-51,'X',53-54,'XXXGY',146,'E',148,'E',150-151,'XV',155 <LIM>  
A:Experimental source: dental enamel  
A:Note: albumin and other serum proteins are also found in bone  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-22/Domain: propeptide #status predicted <PRO>  
F:23-605/Product: serum albumin #status predicted <MAT>  
F:27-199/Domain: serum albumin repeat homology <SA1>  
F:218-391/Domain: serum albumin repeat homology <SA2>  
F:410-589/Domain: serum albumin repeat homology <SA3>  
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,  
F:261/Binding site: bilirubin (lys) #status predicted

Query Match 78.5%; Score 1623.5; DB 1; Length 605;  
Best Local Similarity 75.2%; Pred. No. 2.4e-113;  
Matches 291; Conservative 49; Mismatches 46; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
Db 23 DTYKSEIAHRFNDJGEKFGKGLVIAFSQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 82  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLPLVRPEV 120  
Db 83 NCDKSLHTLFGDKLCAIPSLREHYGDLADCKEKEPERNECFLOHDDNPNLPLVRPEV 141  
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAAPTECCQADKAACLLP 180  
Db 142 VALCADFQEDPKFLGKLYIEIARRHPYFYAPPELLYAIYKDFVSECCQADKAACLLP 201  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEYSKLVTDLT 240  
Db 202 KIEHLREKVLTSAAKQRLKASIQKFGERAFKAWAVARLSORFPKADFTESKIIVTDLAK 261  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 262 VHEKCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 321  
QY 301 DLPSLAADPVESKDVCKNVAEAKDVLGMLFYIYARRHPDYSVLLRLRAKYETTTLEKC 360  
Db 322 DLPNLEHDFVEDKVCNKYKAKADVFLGTFLLYESSRRHPDYSVLLRLRIAKTYEATLEDC 381

QY 361 CAADPHCEYAKVDFDEKPLVBPQNL 387  
Db 382 CAKEDPPACVAFVDFQFTPLVDEPKNL 408

RESULT 6  
ABOS  
serum albumin precursor [validated] - bovine  
N:Alternate names: 67K protein; preproalbumin  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 24-Apr-1984 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000

F;166-174/Product: kinetensin #status experimental <KIP>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,148-125,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 2068; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.8e-146; Indels 0; Gaps 0;  
Matches 387; Conservative 0; Mismatches 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEFHDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEFHDHVKLVNEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLRPEV 144

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 264

QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK 384

QY 361 CAADPHCEYAKVDFEFPKPLVEBPQN 387  
DB 385 CAADPHCEYAKVDFEFPKPLVEBPQN 411

RESULT 2  
A47391  
Serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
F:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding site  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote  
A:Accession: A47391  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295  
A:Note: sequence extracted from NCBI backbone (NCBI:128280, NCBI:128281)  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:21-194/Domain: serum albumin repeat homology <SA1>  
F:213-386/Domain: serum albumin repeat homology <SA2>  
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 96.1%; Score 1988; DB 2; Length 600;  
Best Local Similarity 94.6%; Pred. No. 1.6e-140;  
Matches 366; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEFHDHVKLVNEVTEFAKTCVADESAAE 60  
DB 17 DTHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEFHDHVKLVNEVTEFAKTCVADESAAE 76

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLRPEV 120

DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLRPEV 136

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 180

DB 137 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 196

QY 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240

DB 197 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 256

QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 300

DB 257 VHTCCCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 316

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK 360

DB 317 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK 376

QY 361 CAADPHCEYAKVDFEFPKPLVEBPQN 387

DB 377 CAADPHCEYAKVDFEFPKPLVEBPQN 403

RESULT 3  
S57632  
Serum albumin precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C>Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: J04660; S57632  
F:Wilgier, C.; Grigioni, F.; Hentges, F.  
Gene 169, 235-296, 1996  
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A:Reference number: J04660; MUID:96194824; PMID:8647469  
A:Accession: J04660  
A:Molecule type: mRNA  
A:Residues: 1-608 <HI2>  
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
A:Experimental source: liver  
C:Comment: This protein is the major protein component in plasma. It functions as a mu  
ein has 35 conserved cysteine residues  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: liver; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRP>  
F:125-608/Product: serum albumin #status predicted <MAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 83.5%; Score 1727; DB 2; Length 608;  
Best Local Similarity 80.6%; Pred. No. 4.6e-121;  
Matches 312; Conservative 34; Mismatches 41; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEFHDHVKLVNEVTEFAKTCVADESAAE 60

DB 25 EAHQSEIAHRFNDLGEHFRGLVLFVAFSQYLQCCPEFHDHVKLVNEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLRPEV 120

DB 85 NCEKSLHLLGDKLCTVASLRDKYGEADCCCKEKEPERNECFLOHKDDNPNLPRVLRPEV 144

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 180

DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAFAFTECCQAADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240

DB 205 KYDALREKVLASSAKERLKCASLQKFGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 264

QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 300

DB 265 IHKECCHGDLLECCADRADLAKYICENODSISKLKECKEPLLEKSHCIAEVENDEMPA 324

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A;Reference number: G08292  
A;Accession: G01747  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-120, 'G', 122-455 <MEN>  
A;Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
A;Reference number: S55314; MUID:95275251; PMID:7755581  
A;Accession: S55314  
A;Molecule type: protein  
A;Residues: 19-27 <LED>  
R;Meloun, B.; Moravek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A;Title: Complete amino acid sequence of human serum albumin.  
A;Reference number: A91420; MUID:76187907; PMID:1225573  
A;Accession: A91420  
A;Molecule type: protein  
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
R;Roehr, U.; Spitteller, G.; Tripiel, D.  
Jusztus Liebig's Ann. Chem. 9, 861-884, 1988  
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur  
A;Reference number: S06422  
A;Note: this paper is in German, with an English abstract  
A;Accession: S06422  
A;Molecule type: protein  
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A;Title: Mass spectrometric identification of modifications to human serum albumin treat  
A;Reference number: S36882; MUID:93384321; PMID:8373198  
A;Accession: S36882  
A;Molecule type: protein  
R;Kausler, E.; Spitteller, G.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A;Title: Mass spectrometric identification of modifications to human serum albumin treat  
A;Reference number: S17599; MUID:92126241; PMID:1772598  
A;Accession: S17599  
A;Molecule type: protein  
A;Residues: 25-54/354-357; 431-447 <KAU>  
A;Note: 49-Leu was also found  
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1694, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: A45800  
A;Molecule type: protein  
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
A;Reference number: A03239; MUID:86242180; PMID:3087352  
A;Accession: A03239  
A;Molecule type: protein  
A;Residues: 166-173, 'L' <MOG>  
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A;Title: Mutations in genetic variants of human serum albumin found in Italy.  
A;Reference number: A38255; MUID:91062352; PMID:2247440  
A;Accession: A38255  
A;Molecule type: protein  
A;Residues: 82-105, 'K', 107-110 <GAL2>  
A;Note: this variant is designated albumin Vibo Valentia  
A;Accession: A38255  
A;Molecule type: protein

A;Residues: 76-83, 'K', 85-106 <GAL3>  
A;Note: this variant is designated albumin Torino  
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A;Title: The structural characterization and bilirubin-binding properties of albumin He  
A;Reference number: S33298; MUID:93292504; PMID:8513793  
A;Accession: S33298  
A;Molecule type: protein  
A;Residues: 255-263, 'B', 265-281 <MIN1>  
R;Minchiotti, L.; Galliano, M.; Stopponi, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta  
Biochim. Biophys. Acta 1119, 239-238, 1992  
A;Title: Two albumins with identical electrophoretic mobility are produced by diff  
A;Reference number: S21078; MUID:92190239; PMID:1347703  
A;Accession: S21078  
A;Molecule type: protein  
A;Residues: 354-356, 'K', 358-378 <MIN2>  
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported  
R;He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A;Title: Atomic structure and chemistry of human serum albumin.  
A;Reference number: A46756; MUID:92334427; PMID:1630489  
A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.  
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-41  
A;Reference number: A94442  
A;Contents: annotation; three-dimensional structure and disulfide bonds  
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A;Title: Disulfide bonds in human serum albumin.  
A;Reference number: A90930  
A;Contents: annotation; disulfide bonds  
R;Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A;Reference number: A90299; MUID:78186630; PMID:656055  
A;Contents: annotation; bilirubin-binding site  
R;Peters, T.; Reed, R.G.  
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20  
A;Title: Serum albumin: conformation and active sites.  
A;Reference number: A94408  
A;Contents: annotation; binding sites  
R;Harper, M.E.; Dugaiczky, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene  
A;Reference number: A90028; MUID:8327982; PMID:6192711  
A;Contents: annotation; gene position  
R;Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic aci  
A;Reference number: A46755; MUID:76257808; PMID:955075  
A;Contents: annotation  
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic aci  
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 266-269, 1992  
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosp  
A;Reference number: A56294; MUID:92183981; PMID:1544460  
A;Contents: annotation  
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in  
A;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weat  
C;Comment: A large number of variants of human serum albumin have been described.  
C;Genetics:  
A;Gene: GDB:ALB  
A;Cross-references: GDB:118990; OMIM:103600  
A;Map position: 4q11-4q13  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-609/Product: serum albumin #status experimental <MPT>  
F;29-202/Domain: serum albumin repeat homology <SAL>

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 47.169 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_1\_387

Perfect score: 2068

Sequence: 1 DAHKEVAHRFDLGENFK.....ECVAKVFDFKPLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: Piri:\*

2: Piri2:\*

3: Piri3:\*

4: Piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Match | Length | DB ID    | Description        |
|------------|--------|-------|--------|----------|--------------------|
| 1          | 2068   | 100.0 | 609    | 1 ABHUS  | serum albumin prec |
| 2          | 1988   | 96.1  | 600    | 2 A47391 | serum albumin prec |
| 3          | 1727   | 83.5  | 608    | 2 S57632 | serum albumin prec |
| 4          | 1660.5 | 80.3  | 607    | 1 ABHOS  | serum albumin prec |
| 5          | 1623.5 | 78.5  | 605    | 1 ABPGS  | serum albumin prec |
| 6          | 1618.5 | 78.3  | 607    | 1 ABPOS  | serum albumin prec |
| 7          | 1613.5 | 78.0  | 607    | 1 ABRS   | serum albumin prec |
| 8          | 1610   | 77.9  | 608    | 1 ABRS   | serum albumin prec |
| 9          | 1585   | 76.6  | 609    | 2 JCS838 | albumin - Mongolia |
| 10         | 1264   | 61.1  | 453    | 2 A05139 | serum albumin - mo |
| 11         | 1038.5 | 50.2  | 615    | 1 ABCHS  | serum albumin prec |
| 12         | 898    | 43.4  | 265    | 2 I46986 | albumin - dog (fra |
| 13         | 766    | 37.0  | 607    | 1 ABXL72 | 74K albumin precu  |
| 14         | 737.5  | 35.7  | 608    | 1 ABXL68 | 68K serum albumin  |
| 15         | 724.5  | 35.0  | 609    | 2 JC4358 | alpha-fetoprotein  |
| 16         | 721.5  | 34.9  | 609    | 1 PFPU   | alpha-fetoprotein  |
| 17         | 720.5  | 34.8  | 609    | 1 PFPU   | alpha-fetoprotein  |
| 18         | 691    | 33.4  | 609    | 1 A54906 | afamin precursor - |
| 19         | 656    | 31.7  | 608    | 2 A53195 | alpha-fetoprotein  |
| 20         | 616    | 29.8  | 605    | 1 PPM5   | alpha-fetoprotein  |
| 21         | 603    | 29.2  | 611    | 1 FPRT   | serum albumin prec |
| 22         | 572.5  | 27.7  | 614    | 2 S59317 | serum albumin prec |
| 23         | 468    | 22.6  | 608    | 1 ABONS1 | serum albumin 1 pr |
| 24         | 465    | 22.5  | 608    | 1 ABONS2 | serum albumin 2 pr |
| 25         | 357    | 17.3  | 474    | 1 VYHJ   | vitamin D-binding  |
| 26         | 353    | 17.1  | 382    | 2 A37253 | serum albumin - bu |
| 27         | 352    | 17.0  | 476    | 1 VYRJD  | vitamin D-binding  |
| 28         | 341    | 16.5  | 472    | 1 A35327 | vitamin D-binding  |
| 29         | 311    | 15.0  | 1423   | 1 S27941 | serum albumin - se |

#### RESULT 1

##### ABHUS

serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

N:Contains: kinetensin

C:Species: Homo sapiens (man)

C:Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000

C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3

R:Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houch, C.M.; Najarian, R.C.; Seeb

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K', 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:928591; PIDN:CAA

R:Dugaiczky, A.; Law, S.W.; Demison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936

A:Molecule type: mRNA

A:Residues: 1-120, 'G', 122-609 <DUG>

A:Cross-references: EMBL:V00494; NID:928599; PIDN:CAA23753.1; PID:928590

R:Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and

A:Reference number: I39427; MUID:86140099; PMID:2419329

A:Accession: I39427

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <URA>

A:Cross-references: GB:M13075; NID:9178330; PIDN:AAAS1688.1; PID:9546033

R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family

A:Reference number: I59286; MUID:94181575; PMID:8134387

A:Accession: I59286

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 282-290, 'KSRFDLO' <WAT>

A:Cross-references: GB:S69192; NID:9546032; PIDN:AAAS1688.1; PID:9546033

A>Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia

R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl

A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 589-590, 'ALPRVKNLLQVLP' <WAD>

A:Cross-references: GB:S70799; NID:9547231; PIDN:AAAS1177.1; PID:9547232

A:Cross-references: GB:S70799; NID:9547231; PIDN:AAAS1177.1; PID:9547232

```

RESULT 15
US-10-153-064-125
; Sequence 125, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 125
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-125

Query Match      100.0%; Score 3103; DB 4; Length 677;
Best Local Similarity 100.0%; Pred. No. 7.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 93 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 152
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
Db 153 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 212
Qy 121 DVNCTAFHDNEEFLLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLIP 180
Db 213 DVNCTAFHDNEEFLLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLIP 272
Qy 181 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSGLVTDLT 240
Db 273 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSGLVTDLT 332
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
Db 333 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 392
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 393 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 452
Qy 361 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Db 453 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 512
Qy 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 513 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 572
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
Db 573 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 632
Qy 541 KEQLKAVMDDFAAVFVEKCKKADDDKCTCFEEGKKLVAAASQAALGL 585
Db 633 KEQLKAVMDDFAAVFVEKCKKADDDKCTCFEEGKKLVAAASQAALGL 677

```

Search completed: April 19, 2004, 12:05:17  
Job time : 90.6676 secs



QY 361 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
DB 432 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 491  
QY 421 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 492 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 551  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540  
DB 552 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 511  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 612 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 656

RESULT 13  
US-10-153-064-127  
; Sequence 127, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 127  
; LENGTH: 676  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-064-127

Query Match 100.0%; Score 3103; DB 4; Length 676;  
Best Local Similarity 100.0%; Pred. No. 7.3e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 92 DAHSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 151  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180  
DB 212 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 271  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240  
DB 272 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 331  
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMPA 300  
DB 332 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMPA 391  
QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 392 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 451  
QY 361 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
DB 452 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 511  
QY 421 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 512 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 612 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 656

QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540  
DB 572 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 631  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 632 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 676

RESULT 14  
US-10-153-064-129  
; Sequence 129, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 129  
; LENGTH: 676  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-064-129

Query Match 100.0%; Score 3103; DB 4; Length 676;  
Best Local Similarity 100.0%; Pred. No. 7.3e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 92 DAHSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 151  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180  
DB 212 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 271  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240  
DB 272 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 331  
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMPA 300  
DB 332 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMPA 391  
QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 392 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 451  
QY 361 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
DB 452 CAADPHCEYAKVDFDEPKPLVEBPQNLKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 511  
QY 421 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 512 PTLVEVSRNLGKVGSKCKCKPEAKRMPKABDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540  
DB 572 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 631  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 632 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 676

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 188 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 247  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 248 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 307  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 308 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 367  
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 368 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 427  
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 420  
DB 428 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 487  
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480  
DB 488 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 547  
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 548 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 607  
QY 541 KEOLKAVMDFFAFAVEKCKKADDKETCFABEGKLVAAASQAALGL 585  
DB 608 KEOLKAVMDFFAFAVEKCKKADDKETCFABEGKLVAAASQAALGL 652

## RESULT 11

US-10-153-064-131  
; Sequence 131, Application US/10153064  
; Patent No. 6663485

; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 131

; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-153-064-131

Query Match 100.0%; Score 3103; DB 4; Length 653;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 69 DAHKSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 128  
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120  
DB 129 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 188  
QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 189 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 248  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 249 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 308

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 309 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 368  
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 369 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 428  
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 420  
DB 429 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 488  
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480  
DB 489 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 548  
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 549 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 608  
QY 541 KEOLKAVMDFFAFAVEKCKKADDKETCFABEGKLVAAASQAALGL 585  
DB 609 KEOLKAVMDFFAFAVEKCKKADDKETCFABEGKLVAAASQAALGL 653

## RESULT 12

US-10-153-064-130  
; Sequence 130, Application US/10153064  
; Patent No. 6663485

; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 130

; LENGTH: 656  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-153-064-130

Query Match 100.0%; Score 3103; DB 4; Length 656;  
Best Local Similarity 100.0%; Pred. No. 7e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 72 DAHKSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 131  
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120  
DB 132 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 191  
QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 192 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 251  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 252 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 311  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 312 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 371  
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 372 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 431

; MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2;  
US-09-984-186-2

Query Match 100.0%; Score 3103; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.3e-287;  
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGEENFKALVLIAPFAQYLQQCPFEDHVKL VNEVTEFAKTCVADSAE 60  
DB 25 DAKHSEVAHRFKDGLGEENFKALVLIAPFAQYLQQCPFEDHVKL VNEVTEFAKTCVADSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNPNLPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNPNLPRLVRPEV 144  
QY 121 DVMTAFPHNBEETFLKKLYIEIARRHPFYFPAPELLFFAKRYKAAPTCECQAADKAAACLLP 180  
DB 145 DVMTAFPHNBEETFLKKLYIEIARRHPFYFPAPELLFFAKRYKAAPTCECQAADKAAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 240  
DB 205 KLDELDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 264  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCNKYAERAKDVFLGMFLYEYARRHPDYSVLLLR LAKTYETTILEKC 360  
DB 325 DLPSLAADFVESKDVCNKYAERAKDVFLGMFLYEYARRHPDYSVLLLR LAKTYETTILEKC 384  
QY 361 CAAADPHCYAKVFDFEPLVEEQNLIKQNCELFEQLGEYKFNALLVRYTKVPQVST 420  
DB 385 CAAADPHCYAKVFDFEPLVEEQNLIKQNCELFEQLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKHPKAMPKPCAEVDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRLNGKVGSKCKHPKAMPKPCAEVDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540  
DB 505 LVNRRCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 564  
QY 541 KEOLKAVMDDFAAFVKECKCKADDKETCFEAEGKKLVAASAQALGL 585  
DB 565 KEOLKAVMDDFAAFVKECKCKADDKETCFEAEGKKLVAASAQALGL 609

RESULT 9  
US-10-153-064-133  
Sequence 133, Application US/10153064  
Patent No. 6663485  
GENERAL INFORMATION:  
APPLICANT: Bell et al.  
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
FILE REFERENCE: PF556  
CURRENT APPLICATION NUMBER: US/10/153,064  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 60/293,212  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 133  
LENGTH: 651  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-153-064-133

Query Match 100.0%; Score 3103; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGEENFKALVLIAPFAQYLQQCPFEDHVKL VNEVTEFAKTCVADSAE 60  
DB 68 DAKHSEVAHRFKDGLGEENFKALVLIAPFAQYLQQCPFEDHVKL VNEVTEFAKTCVADSAE 127  
QY 61 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNPNLPRLVRPEV 120  
DB 128 NCDKSLHTLFGDKLCTVATLRITYGEMADCCAKQEPERNECFLOHKDNPNLPRLVRPEV 187

```
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fleer, Reinhard
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHAFKDLGEENFKALVLIAPAOYLQCCPEPDHVKLVNEVTEPAKTCVADESAAE 60
Db 25 DAHKEVAHAFKDLGEENFKALVLIAPAOYLQCCPEPDHVKLVNEVTEPAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPPKAFPAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPPKAFPAEVSKLVTDLTK 264
Qy 241 VHTCECHGDLLECADRADLAKVICENODSISKLKECCPKLLEKSHCIATVENDMPA 300
Db 265 VHTCECHGDLLECADRADLAKVICENODSISKLKECCPKLLEKSHCIATVENDMPA 324

; RESULT 8
; US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
QY 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 585

## RESULT 5

US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Patent No. 6663485  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 3103; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.2e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 25 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 144  
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELREGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFPKAEFVSKLVTDLT 240  
DB 205 KLDELREGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFPKAEFVSKLVTDLT 264  
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDMPA 300  
DB 265 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDMPA 324  
QY 301 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYEARHPDYSVVLRLAKTYETTLEK 360  
DB 325 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYEARHPDYSVVLRLAKTYETTLEK 384  
QY 361 CAADDPHECYAKVDFEKPILVEEPQNLIKQNCLEFQELGGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADDPHECYAKVDFEKPILVEEPQNLIKQNCLEFQELGGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKKPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTCKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTCKCTES 504  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
DB 505 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 564  
QY 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 609

## RESULT 7

US-08-797-689-2  
; Sequence 2, Application US/08797689

## RESULT 6

US-09-976-594-977  
; Sequence 77, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furtress, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 977  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1  
US-09-976-594-977

Query Match 100.0%; Score 3103; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.2e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 25 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 144  
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELREGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFPKAEFVSKLVTDLT 240  
DB 205 KLDELREGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFPKAEFVSKLVTDLT 264  
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDMPA 300  
DB 265 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDMPA 324  
QY 301 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYEARHPDYSVVLRLAKTYETTLEK 360  
DB 325 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYEARHPDYSVVLRLAKTYETTLEK 384  
QY 361 CAADDPHECYAKVDFEKPILVEEPQNLIKQNCLEFQELGGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADDPHECYAKVDFEKPILVEEPQNLIKQNCLEFQELGGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKKPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTCKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTCKCTES 504  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
DB 505 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 564  
QY 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEEGKKLVAASQAALGL 609

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
 Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585

## RESULT 3

US-08-769-746-2  
 ; Sequence 2, Application US/08769746  
 ; Patent No. 6274305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sonnenschein, Carlos  
 ; APPLICANT: Soco, Ana M.  
 ; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Medlen & Carroll, LLP  
 ; STREET: 220 Montgomery Street, Suite 2200  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States of America  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/769,746  
 ; FILING DATE: 19-DEC-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Carroll, Peter G.  
 ; REGISTRATION NUMBER: 32,837  
 ; REFERENCE/DOCKET NUMBER: MBRI-02584  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 705-8410  
 ; TELEFAX: (415) 397-8338  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 585 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-287;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGKLVIAFAQYLOCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHKSEVAHRFKDLGKLVIAFAQYLOCCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
 Db 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

QY 181 KLDELDEBKASSAKQRLKASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240  
 Db 181 KLDELDEBKASSAKQRLKASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240

QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 Db 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVAREHPDYSVVLRLAKTYETTLK 360  
 Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVAREHPDYSVVLRLAKTYETTLK 360

QY 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480  
 Db 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480

QY 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELPQOLGEYKFNALLVRYTKVQPVST 420  
 Db 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELPQOLGEYKFNALLVRYTKVQPVST 420

QY 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480  
 Db 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540  
 Db 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
 Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585

## RESULT 4

US-10-153-064-5  
 ; Sequence 5, Application US/10153064  
 ; Patent No. 6663485  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell et al.  
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
 ; FILE REFERENCE: PF556  
 ; CURRENT APPLICATION NUMBER: US/10/153,064  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/293,212  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 ; US-10-153-064-5

Query Match 100.0%; Score 3103; DB 4; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-287;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGKLVIAFAQYLOCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHKSEVAHRFKDLGKLVIAFAQYLOCCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
 Db 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

QY 181 KLDELDEBKASSAKQRLKASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240  
 Db 181 KLDELDEBKASSAKQRLKASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240

QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 Db 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVAREHPDYSVVLRLAKTYETTLK 360  
 Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVAREHPDYSVVLRLAKTYETTLK 360

QY 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELPQOLGEYKFNALLVRYTKVQPVST 420  
 Db 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELPQOLGEYKFNALLVRYTKVQPVST 420

QY 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480  
 Db 421 PTLVEVSRNLGKVGSKCKKHPKARMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: Region
;   LOCATION: 369..419
;   OTHER INFORMATION: /note= "Alternative C-termini of
;   OTHER INFORMATION: HSA(1-n)"
; FEATURE:
;   NAME/KEY: Region
;   LOCATION: 1..585
;   OTHER INFORMATION: /note= "Amino acid sequence of
;   OTHER INFORMATION: natural HSA"
;
US-08-153-799-14

Query Match      100.0%; Score 3103; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVTPKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVTPKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

```

## RESULT 2

```

US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue

```

```

; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-702-572-2

```

```

Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVTPKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVTPKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 89.6676 Seconds

(without alignments)  
336.813 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAHSEVAHFYDGLGENPK.....TCFAEEGKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/2/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/prodata/2/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/prodata/2/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/prodata/2/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/prodata/2/iaa/PCITUS COMB.pep.\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 3103  | 100.0       | 585    | 1  | US-08-153-799-14  |
| 2          | 3103  | 100.0       | 585    | 2  | US-08-702-572-2   |
| 3          | 3103  | 100.0       | 585    | 3  | US-08-769-746-2   |
| 4          | 3103  | 100.0       | 585    | 4  | US-10-153-064-5   |
| 5          | 3103  | 100.0       | 609    | 4  | US-10-153-064-7   |
| 6          | 3103  | 100.0       | 609    | 4  | US-09-576-594-977 |
| 7          | 3103  | 100.0       | 610    | 2  | US-08-797-689-2   |
| 8          | 3103  | 100.0       | 610    | 4  | US-09-984-186-2   |
| 9          | 3103  | 100.0       | 651    | 4  | US-10-153-064-133 |
| 10         | 3103  | 100.0       | 652    | 4  | US-10-153-064-132 |
| 11         | 3103  | 100.0       | 653    | 4  | US-10-153-064-131 |
| 12         | 3103  | 100.0       | 656    | 4  | US-10-153-064-130 |
| 13         | 3103  | 100.0       | 676    | 4  | US-10-153-064-127 |
| 14         | 3103  | 100.0       | 676    | 4  | US-10-153-064-129 |
| 15         | 3103  | 100.0       | 677    | 4  | US-10-153-064-125 |
| 16         | 3103  | 100.0       | 680    | 4  | US-10-153-064-123 |
| 17         | 3103  | 100.0       | 783    | 1  | US-08-256-938-2   |
| 18         | 3103  | 100.0       | 787    | 1  | US-08-566-938-4   |
| 19         | 3103  | 100.0       | 787    | 2  | US-08-797-689-16  |
| 20         | 3103  | 100.0       | 787    | 4  | US-09-984-186-16  |
| 21         | 3099  | 99.9        | 609    | 1  | US-08-222-619-3   |
| 22         | 3099  | 99.9        | 609    | 1  | US-08-433-037-4   |
| 23         | 3099  | 99.9        | 609    | 4  | US-08-897-956A-2  |
| 24         | 3099  | 99.9        | 609    | 5  | PCT-US95-04075-3  |
| 25         | 3095  | 99.7        | 978    | 4  | US-08-897-956A-3  |
| 26         | 3093  | 99.7        | 585    | 1  | US-08-448-196A-3  |
| 27         | 3093  | 99.7        | 585    | 2  | US-08-984-176-1   |

|    |        |      |      |   |                   |
|----|--------|------|------|---|-------------------|
| 28 | 3092.5 | 99.7 | 652  | 4 | US-10-153-064-96  |
| 29 | 3092.5 | 99.7 | 652  | 4 | US-10-153-064-99  |
| 30 | 3092.5 | 99.7 | 652  | 4 | US-10-153-064-105 |
| 31 | 3092.5 | 99.7 | 660  | 4 | US-10-153-064-90  |
| 32 | 3092.5 | 99.7 | 660  | 4 | US-10-153-064-93  |
| 33 | 3092.5 | 99.7 | 676  | 4 | US-10-153-064-95  |
| 34 | 3092.5 | 99.7 | 676  | 4 | US-10-153-064-98  |
| 35 | 3092.5 | 99.7 | 676  | 4 | US-10-153-064-104 |
| 36 | 3092.5 | 99.7 | 684  | 4 | US-10-153-064-92  |
| 37 | 3092.5 | 99.7 | 1184 | 4 | US-10-153-064-89  |
| 38 | 3086.5 | 99.5 | 668  | 4 | US-10-153-064-102 |
| 39 | 3086.5 | 99.5 | 692  | 4 | US-10-153-064-101 |
| 40 | 2458.5 | 79.2 | 583  | 1 | US-08-448-196A-5  |
| 41 | 2450.5 | 79.0 | 583  | 1 | US-08-448-196A-4  |
| 42 | 2432.5 | 78.4 | 583  | 1 | US-08-448-196A-6  |
| 43 | 2426   | 78.2 | 584  | 1 | US-08-448-196A-7  |
| 44 | 2389   | 77.0 | 582  | 1 | US-08-134-638-1   |
| 45 | 1249.5 | 40.3 | 609  | 1 | US-08-222-619-4   |

## ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/153.799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

|                    |
|--------------------|
| Sequence 96, Appl  |
| Sequence 99, Appl  |
| Sequence 105, Appl |
| Sequence 90, Appl  |
| Sequence 93, Appl  |
| Sequence 95, Appl  |
| Sequence 98, Appl  |
| Sequence 104, Appl |
| Sequence 92, Appl  |
| Sequence 89, Appl  |
| Sequence 102, Appl |
| Sequence 101, Appl |
| Sequence 5, Appl   |
| Sequence 4, Appl   |
| Sequence 6, Appl   |
| Sequence 7, Appl   |
| Sequence 1, Appl   |
| Sequence 4, Appl   |



```

; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 15
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VADESAEN 8
Db      54 VADESAEN 61

Search completed: April 19, 2004, 12:54:58
Job time : 4.36842 secs

```

```
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 11
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLyS)
; FILE REFERENCE: Dyx-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 12
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 13
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 14
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; FAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2
Query Match 100.0%; Score 38; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 7
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445
Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 8
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

```

```
; TITLE OF INVENTION: CHIMERIC POLYPEPTIDES OF SERUM ALBUMIN AND USES RELATED THERETO
; FILE REFERENCE: GPCI-P01-109
; CURRENT APPLICATION NUMBER: US/09/768,183
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/764918
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 09/619285
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/144534
; PRIOR FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-768-183-6

Query Match      100.0%; Score 38; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      4 VADESAEN 11

RESULT 3
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-24

Query Match      100.0%; Score 38; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 4
US-10-074-956-27
; Sequence 27, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-27

Query Match      100.0%; Score 38; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

RESULT 5
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-28

Query Match      100.0%; Score 38; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

RESULT 6
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.36842 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61  
Perfect score: 38  
Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 38    | 100.0       | 14     | 12 | US-09-764-918-6   |
| 2          | 38    | 100.0       | 14     | 12 | US-09-768-183-6   |
| 3          | 38    | 100.0       | 195    | 13 | US-10-074-956-24  |
| 4          | 38    | 100.0       | 241    | 13 | US-10-074-956-27  |
| 5          | 38    | 100.0       | 268    | 13 | US-10-074-956-28  |
| 6          | 38    | 100.0       | 585    | 9  | US-09-929-552-2   |
| 7          | 38    | 100.0       | 585    | 10 | US-09-932-613-445 |
| 8          | 38    | 100.0       | 585    | 10 | US-09-984-010-26  |
| 9          | 38    | 100.0       | 585    | 10 | US-09-833-041-18  |
| 10         | 38    | 100.0       | 585    | 10 | US-09-833-117-18  |
| 11         | 38    | 100.0       | 585    | 10 | US-09-932-322-445 |
| 12         | 38    | 100.0       | 585    | 10 | US-09-832-501-18  |
| 13         | 38    | 100.0       | 585    | 11 | US-09-833-118-18  |
| 14         | 38    | 100.0       | 585    | 11 | US-09-833-245-18  |
| 15         | 38    | 100.0       | 585    | 12 | US-10-424-999-11  |

|    |    |       |     |    |                   |                   |
|----|----|-------|-----|----|-------------------|-------------------|
| 16 | 38 | 100.0 | 585 | 12 | US-10-425-000-31  | Sequence 31, Appl |
| 17 | 38 | 100.0 | 585 | 12 | US-10-433-108-34  | Sequence 34, Appl |
| 18 | 38 | 100.0 | 585 | 13 | US-10-153-064-5   | Sequence 5, Appl  |
| 19 | 38 | 100.0 | 585 | 14 | US-10-153-604A-5  | Sequence 5, Appl  |
| 20 | 38 | 100.0 | 585 | 14 | US-10-319-263-1   | Sequence 1, Appl  |
| 21 | 38 | 100.0 | 585 | 14 | US-10-319-263-2   | Sequence 2, Appl  |
| 22 | 38 | 100.0 | 585 | 14 | US-10-414-469-1   | Sequence 1, Appl  |
| 23 | 38 | 100.0 | 585 | 14 | US-10-414-469-2   | Sequence 2, Appl  |
| 24 | 38 | 100.0 | 585 | 14 | US-10-413-831-1   | Sequence 1, Appl  |
| 25 | 38 | 100.0 | 585 | 14 | US-10-413-831-2   | Sequence 2, Appl  |
| 26 | 38 | 100.0 | 585 | 15 | US-10-413-832-1   | Sequence 1, Appl  |
| 27 | 38 | 100.0 | 585 | 15 | US-10-413-832-2   | Sequence 2, Appl  |
| 28 | 38 | 100.0 | 585 | 15 | US-10-414-386-1   | Sequence 1, Appl  |
| 29 | 38 | 100.0 | 585 | 15 | US-10-414-386-2   | Sequence 2, Appl  |
| 30 | 38 | 100.0 | 585 | 15 | US-10-233-675A-11 | Sequence 11, Appl |
| 31 | 38 | 100.0 | 585 | 15 | US-10-462-262-26  | Sequence 26, Appl |
| 32 | 38 | 100.0 | 604 | 10 | US-09-984-010-7   | Sequence 7, Appl  |
| 33 | 38 | 100.0 | 609 | 10 | US-09-919-039-370 | Sequence 370, App |
| 34 | 38 | 100.0 | 609 | 12 | US-10-609-346-12  | Sequence 12, Appl |
| 35 | 38 | 100.0 | 609 | 13 | US-10-153-064-7   | Sequence 7, Appl  |
| 36 | 38 | 100.0 | 609 | 14 | US-10-153-604A-7  | Sequence 7, Appl  |
| 37 | 38 | 100.0 | 609 | 14 | US-10-365-623-23  | Sequence 23, Appl |
| 38 | 38 | 100.0 | 610 | 9  | US-09-984-186-2   | Sequence 2, Appl  |
| 39 | 38 | 100.0 | 610 | 14 | US-10-237-667-2   | Sequence 2, Appl  |
| 40 | 38 | 100.0 | 610 | 14 | US-10-237-708-2   | Sequence 2, Appl  |
| 41 | 38 | 100.0 | 610 | 14 | US-10-237-866-2   | Sequence 2, Appl  |
| 42 | 38 | 100.0 | 610 | 14 | US-10-237-871-2   | Sequence 2, Appl  |
| 43 | 38 | 100.0 | 610 | 14 | US-10-237-871-2   | Sequence 2, Appl  |
| 44 | 38 | 100.0 | 616 | 12 | US-10-433-108-13  | Sequence 13, Appl |
| 45 | 38 | 100.0 | 624 | 12 | US-10-433-108-16  | Sequence 16, Appl |

ALIGNMENTS

RESULT 1  
US-09-764-918-6  
; Sequence 6, Application US/09764918  
; Publication No. US20010056075A1  
; GENERAL INFORMATION:  
; APPLICANT: Gyuris, J.  
; TITLE OF INVENTION: CHIMERIC POLYPEPTIDES OF SERUM ALBUMIN AND USES RELATED THERETO  
; FILE REFERENCE: GPC1-P02-109  
; CURRENT APPLICATION NUMBER: US/09/764,918  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 09/619285  
; PRIOR FILING DATE: 2000-07-19  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-764-918-6

Query Match 100.0%; Score 38; DB 12; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8  
DB 4 VADESAEN 11

RESULT 2  
US-09-768-183-6  
; Sequence 6, Application US/09768183  
; Publication No. US20020048571A1  
; GENERAL INFORMATION:  
; APPLICANT: Gyuris, J.  
; APPLICANT: Lamphere, L.  
; APPLICANT: Morris, A.

PR 29-DEC-2000; 2000US-0258764P.  
 PR 14-JUN-2001; 2001US-0298317P.  
 XX  
 XX PA (ZYCO-) ZYCO INC.  
 XX  
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 PI  
 XX WPI; 2002-195801/25.  
 DR  
 XX  
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 PT stimulating hormone concatamer or its analog, for treating inflammatory  
 PT or autoimmune disorders.  
 XX  
 XX Example 2; Page 48; 89pp; English.  
 PS  
 XX The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 CC hormone (MSH) concatamer. The sequences are useful for treating an  
 CC individual suffering from, or at risk of, a disorder of the immune system  
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 CC present sequence is a peptide described in the exemplification of the  
 CC invention  
 XX  
 SQ Sequence 236 AA;  
 Query Match 100.0%; Score 38; DB 5; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 7.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 Db |||||  
 78 VADESAEN 85  
 Search completed: April 19, 2004, 11:51:16  
 Job time : 5.70545 secs



RESULT 11  
 AAO17048  
 ID AAO17048 standard; protein; 195 AA.  
 XX AC AAO17048;  
 XX DT 29-MAY-2002 (first entry)  
 XX DE Human serum albumin (1-195) SEQ ID NO: 56.  
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;  
 XX KW Alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;  
 XX KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;  
 XX KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
 XX KW antineoplastic; antibacterial; dermatological; antipsoriatic;  
 XX KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;  
 XX KW diabetes; uveitis; coeliac disease.  
 XX OS Homo sapiens.  
 XX PN WO200206316-A2.  
 XX PD 24-JAN-2002.  
 XX PF 16-JUL-2001; 2001WO-US022263.  
 XX PR 14-JUL-2000; 2000US-0218381P.  
 XX PR 18-AUG-2000; 2000US-0226382P.  
 XX PR 06-OCT-2000; 2000US-0238380P.  
 XX PR 29-DEC-2000; 2000US-0258764P.  
 XX PR 14-JUN-2001; 2001US-0298317P.  
 XX PA (ZYCO-) ZYCO INC.  
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;  
 XX WPI; 2002-195801/25.  
 XX PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte  
 XX PT stimulating hormone concatamer or its analog, for treating inflammatory  
 XX PT or autoimmune disorders.  
 XX PS Example 2; Page 46; 89pp; English.  
 XX CC The present invention relates to a nucleic acid comprising a sequence  
 XX CC encoding a fusion polypeptide having an alpha-melanocyte stimulating  
 XX CC hormone (MSH) concatamer. The sequences are useful for treating an  
 XX CC individual suffering from, or at risk of, a disorder of the immune system  
 XX CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid  
 XX CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact  
 XX CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,  
 XX CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The  
 XX CC present sequence is a protein described in the exemplification of the  
 XX CC invention  
 XX SQ Sequence 195 AA;  
 Query Match 100.0%; Score 38; DB 5; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 Db 54 VADESAEN 61  
 RESULT 12  
 ABU10022  
 ID ABU10022 standard; protein; 195 AA.  
 XX AC ABU10022;  
 XX DT 28-JUL-2000 (first entry)

DT 31-JUL-2003 (first entry)  
 XX DE Human serum albumin residues 1-195.  
 XX KW Bladder disorder; cytostatic; antiinflammatory; immune response;  
 XX KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;  
 XX KW bladder cancer; tumour; interstitial cystitis; inflammation;  
 XX KW alpha-MSH concatamer; melanocyte stimulating hormone; human;  
 XX KW serum albumin.  
 XX OS Homo sapiens.  
 XX PN US2002193332-A1.  
 XX PD 19-DEC-2002.  
 XX PF 12-FEB-2002; 2002US-00074956.  
 XX PR 12-FEB-2001; 2001US-0268175P.  
 XX PA (HEDL/) HEDLEY M L.  
 XX PI Hedley ML;  
 XX WPI; 2003-447327/42.  
 XX PT Modulating immune responses in a mammal with a bladder disorder e.g.  
 XX PT bladder cancer, by administering nucleic acids comprising un-methylated  
 XX PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to  
 XX PT the mammal.  
 XX PS Example 2; Page 9; 17pp; English.  
 XX CC The invention describes a method of modulating an immune response in a  
 XX CC mammal, comprising identifying a mammal that has or is at risk for having  
 XX CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)  
 XX CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated  
 XX CC nucleic acid (N2) comprising a sequence encoding alpha-MSH to the mammal;  
 XX CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The  
 XX CC method is useful for modulating immune response in a mammal having a  
 XX CC bladder disorder, where administration of (N1) results in an amelioration  
 XX CC of one or more symptoms of the disorder. Preferably, the method is useful  
 XX CC for modulating immune response in a mammal having bladder cancer (where  
 XX CC administration of (N1) results in a decrease in tumour size or activity),  
 XX CC or for modulating immune response in a mammal having interstitial  
 XX CC cystitis (where administration of (N1) results in a modulation of the  
 XX CC immune response from Th2 response to a Th1 response). The method is also  
 XX CC useful for modulating immune response in a mammal having bladder disorder  
 XX CC that is characterised by inflammation which is associated with symptoms  
 XX CC of interstitial cystitis or associated with a disruption of the integrity  
 XX CC of the bladder lining. This is the amino acid sequence of human serum  
 XX CC albumin residues 1-195 that can be used in the creation of melanocyte  
 XX CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the  
 XX CC fusion protein when expressed in mammalian cells  
 XX SQ Sequence 195 AA;  
 Query Match 100.0%; Score 38; DB 7; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 Db 54 VADESAEN 61  
 RESULT 13  
 AAY83947  
 ID AAY83947 standard; protein; 204 AA.  
 XX AC AAY83947;  
 XX DT 28-JUL-2000 (first entry)



CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX  
SQ Sequence 137 AA;

Query Match 100.0%; Score 38; DB 4; Length 137;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VADESAEN 8  
| | | | |  
Db 83 VADESAEN 90

RESULT 9  
AAU33073  
ID AAU33073 standard; protein; 156 AA.  
XX  
AC AAU33073;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3564.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR 26-JAN-2001; 2001US-00770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
PS Claim 20; Page 704; 765pp; English.  
XX

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 156 AA;

OY 1 VADESAEN 8  
| | | | |  
Db 83 VADESAEN 90

Query Match 100.0%; Score 38; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VADESAEN 8  
| | | | |  
Db 9 VADESAEN 16

RESULT 10  
AAU33271  
ID AAU33271 standard; protein; 156 AA.  
XX  
AC AAU33271;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3762.  
XX

Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US008656.  
XX  
PR 18-APR-2000; 2000US-00552929.  
PR 26-JAN-2001; 2001US-00770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
XX  
PS Claim 20; Page 751; 765pp; English.  
XX

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 156 AA;

Query Match 100.0%; Score 38; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VADESAEN 8  
| | | | |  
Db 9 VADESAEN 16

PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-514838/56.  
 DR N-PSDB; AA182560.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX Claim 20; SEQ ID NO 16521; 1399pp + Sequence Listing; English.  
 PS  
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 127 AA;  
 Query Match 100.0%; Score 38; DB 4; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 |||||  
 Db 38 VADESAEN 45  
 RESULT 7  
 AAO09589  
 ID AAO09589 standard; protein; 132 AA.  
 AC AAO09589;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polypeptide SEQ ID NO 23481.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 XX Homo sapiens.  
 OS  
 XX WO200164835-A2.  
 PN  
 XX 07-SEP-2001.  
 PD  
 XX 26-FEB-2001; 2001WO-US004927.  
 PF  
 XX 28-FEB-2000; 2000US-00515126.  
 PR  
 XX 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-514838/56.  
 DR N-PSDB; AA189520.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX

PS Claim 20; SEQ ID NO 23481; 1399pp + Sequence Listing; English.  
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 132 AA;  
 Query Match 100.0%; Score 38; DB 4; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VADESAEN 8  
 |||||  
 Db 82 VADESAEN 89  
 RESULT 8  
 AAO12088  
 ID AAO12088 standard; protein; 137 AA.  
 AC AAO12088;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polypeptide SEQ ID NO 25980.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 XX Homo sapiens.  
 OS  
 XX WO200164835-A2.  
 PN  
 XX 07-SEP-2001.  
 PD  
 XX 26-FEB-2001; 2001WO-US004927.  
 PF  
 XX 28-FEB-2000; 2000US-00515126.  
 PR  
 XX 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-514838/56.  
 DR N-PSDB; AA192019.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX Claim 20; SEQ ID NO 25980; 1399pp + Sequence Listing; English.  
 PS  
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

```

DT 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 18339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI84378.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 18339; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and/or
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 120 AA;
XX
XX Query Match 100.0%; Score 38; DB 4; Length 120;
XX Best Local Similarity 100.0%; Pred. No. 3.4;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VADESAEN 8
XX |||||
XX 18 VADESAEN 25
XX
XX Db
XX
XX RESULT 5
XX AAO04415
XX ID AAO04415 standard; protein; 124 AA.
XX
XX AC AAO04415;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 18307.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
OS

```

---

```

XX PN WO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI84346.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 18307; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and/or
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 124 AA;
XX
XX Query Match 100.0%; Score 38; DB 4; Length 124;
XX Best Local Similarity 100.0%; Pred. No. 3.6;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VADESAEN 8
XX |||||
XX 76 VADESAEN 83
XX
XX Db
XX
XX RESULT 6
XX AAO02629
XX ID AAO02629 standard; protein; 127 AA.
XX
XX AC AAO02629;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 16521.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX
XX

```

```

CC cell death in an organism. The present sequence is a fragment of a
CC modified murine serum albumin protein
XX
SQ Sequence 14 AA;

  Query Match      100.0%; Score 38; DB 6; Length 14;
  Best Local Similarity 100.0%; Pred. No. 0.32;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
   |||||
DB 4 VADESAEN 11

RESULT 2
AAO02642
ID AAO02642 standard; protein; 116 AA.
XX
AC AAO02642;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 16534.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX N-PSDB; AAI82573.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 16534; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 116 AA;

  Query Match      100.0%; Score 38; DB 4; Length 116;
  Best Local Similarity 100.0%; Pred. No. 3.3;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
   |||||
DB 4 VADESAEN 11

RESULT 3
AAO04446
ID AAO04446 standard; protein; 118 AA.
XX
AC AAO04446;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 18338.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX N-PSDB; AAI84377.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 18338; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 118 AA;

  Query Match      100.0%; Score 38; DB 4; Length 118;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
   |||||
DB 76 VADESAEN 83

RESULT 4
AAO04447
ID AAO04447 standard; protein; 120 AA.
XX
AC AAO04447;
XX

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.70545 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_54\_61

Perfect score: 38

Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 38    | 100.0       | 14     | AAO19555 | AAO19555 Modified  |
| 2          | 38    | 100.0       | 116    | AAO02642 | AAO02642 Human pol |
| 3          | 38    | 100.0       | 118    | AAO04446 | AAO04446 Human pol |
| 4          | 38    | 100.0       | 120    | AAO04447 | AAO04447 Human pol |
| 5          | 38    | 100.0       | 124    | AAO04415 | AAO04415 Human pol |
| 6          | 38    | 100.0       | 127    | AAO02629 | AAO02629 Human pol |
| 7          | 38    | 100.0       | 132    | AAO09589 | AAO09589 Human pol |
| 8          | 38    | 100.0       | 137    | AAO12088 | AAO12088 Human pol |
| 9          | 38    | 100.0       | 156    | AAU33073 | AAU33073 Novel hum |
| 10         | 38    | 100.0       | 156    | AAU33271 | AAU33271 Novel hum |
| 11         | 38    | 100.0       | 195    | AAO17048 | AAO17048 Human ser |
| 12         | 38    | 100.0       | 195    | ABU10022 | ABU10022 Human ser |
| 13         | 38    | 100.0       | 204    | AAU83947 | AAU83947 Yeast cod |
| 14         | 38    | 100.0       | 214    | AAU29874 | AAU29874 Novel hum |
| 15         | 38    | 100.0       | 236    | AAO17051 | AAO17051 Human alb |
| 16         | 38    | 100.0       | 241    | AAO16984 | AAO16984 Alpha-MSH |
| 17         | 38    | 100.0       | 241    | ABU10025 | ABU10025 Alpha-MSH |
| 18         | 38    | 100.0       | 242    | AAO16985 | AAO16985 Alpha-MSH |
| 19         | 38    | 100.0       | 244    | AAO16986 | AAO16986 Alpha-MSH |
| 20         | 38    | 100.0       | 245    | AAO16987 | AAO16987 Alpha-MSH |
| 21         | 38    | 100.0       | 245    | AAO16988 | AAO16988 Alpha-MSH |
| 22         | 38    | 100.0       | 268    | AAO16989 | AAO16989 Alpha-MSH |
| 23         | 38    | 100.0       | 268    | ABU10026 | ABU10026 Alpha-MSH |
| 24         | 38    | 100.0       | 289    | AAU29581 | AAU29581 Novel hum |
| 25         | 38    | 100.0       | 289    | AAU29575 | AAU29575 Novel hum |

|    |    |       |     |   |          |           |
|----|----|-------|-----|---|----------|-----------|
| 26 | 38 | 100.0 | 303 | 2 | AA14178  | Human ser |
| 27 | 38 | 100.0 | 373 | 1 | AA090387 | N-termina |
| 28 | 38 | 100.0 | 388 | 1 | AA090389 | N-termina |
| 29 | 38 | 100.0 | 389 | 1 | AA090390 | N-termina |
| 30 | 38 | 100.0 | 390 | 1 | AA090391 | N-termina |
| 31 | 38 | 100.0 | 401 | 4 | AAU29876 | Novel hum |
| 32 | 38 | 100.0 | 407 | 1 | AA090392 | N-termina |
| 33 | 38 | 100.0 | 500 | 7 | ADD32019 | Heterolog |
| 34 | 38 | 100.0 | 584 | 6 | ABG72381 | Mature hu |
| 35 | 38 | 100.0 | 585 | 1 | AA093344 | Sequence  |
| 36 | 38 | 100.0 | 585 | 1 | AA090388 | Mature hu |
| 37 | 38 | 100.0 | 585 | 1 | AA091422 | Human nor |
| 38 | 38 | 100.0 | 585 | 2 | AA095318 | Human ser |
| 39 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 40 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 41 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 42 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 43 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 44 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |
| 45 | 38 | 100.0 | 585 | 2 | AA08457  | Human ser |

ALIGNMENTS

RESULT 1

AAO19555

ID AAO19555 standard; peptide; 14 AA.

XX

AC AAO19555;

XX

DT 28-JAN-2003 (first entry)

XX

DE Modified mouse serum albumin fragment #1.

XX

KW Serum albumin; chimeric peptide; angiogenesis inhibition; cancer;

XX

XX apoptosis; cytostatic.

OS Mus sp.

OS Synthetic.

XX

FH Key

FT Cross-links

FT Location/Qualifiers

FT 3..12

FT /note= "cyclise the molecule"

XX

XX WO200270549-A2.

XX

PD 12-SEP-2002.

XX

XX 18-JAN-2002; 2002WO-US001256.

XX

XX 18-JAN-2001; 2001US-00764918.

XX

XX 23-JAN-2001; 2001US-00768183.

XX

XX (GPCB-) GPC BIOTECH INC.

XX

XX Gyuris J, Lamphere L, Morris A;

XX

XX WPI; 2003-018690/01.

XX

XX New chimeric polypeptide comprising a serum albumin protein having a

XX biologically active heterologous peptide sequence inserted into it,

XX useful for modulating cell proliferation, cell differentiation, or cell

XX death in an organism.

XX

XX Disclosure; Fig 5; 69pp; English.

XX

XX The present invention relates to chimeric peptides comprising a serum

XX albumin protein having a biologically active heterologous peptide

XX sequence inserted into it. The chimeric peptide exhibits increased

XX biological activity relative to the heterologous peptide sequence itself.

XX The chimeric polypeptide may be used for treating a disease in an

XX organism, or for modulating cell proliferation, cell differentiation, or

XX

CC three synthetic fragments (AA10092-A10094) joined by recombinant DNA  
CC technology. Each HSA fragment was synthesised from overlapping  
CC oligonucleotide fragments that were extended. This sequence represents  
CC the complete sequence of the HSA encoded by the human gene with a yeast  
CC codon bias. The invention also covers a recombinant expression vector,  
CC yeast host cells carrying the recombinant expression vector and the  
CC process for producing human serum albumin in the yeast host cell,  
CC especially in secretory mode  
XX Sequence 585 AA;  
SQ  
Query Match 100.0%; Score 2068; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOQCPEFHVKLNVETFAKTCVADESAE 60  
DB 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOQCPEFHVKLNVETFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTTLEKC 360  
QY 361 CAADPHECYAKVDFDEPKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFDEPKPLVEEPQNL 387  
RESULT 15  
AAMS2567  
ID AAMS2567 standard; protein; 585 AA.  
XX  
AC AAMS2567;  
XX  
DT 05-FEB-2002 (first entry)  
XX  
DE Mature human serum albumin.  
XX  
KW Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiatic;  
KW neotropic; neuroprotective; gene therapy; immune disorder; wound healing;  
KW hyperproliferative disorder; renal disorder; cardiovascular disorder;  
KW respiratory disorder; neurological disease; endocrine disorder;  
KW reproductive system disorder; infectious disease;  
KW gastrointestinal disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200179444-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US012013.  
XX  
PR 12-APR-2000; 2000US-0229359P.  
XX  
PR 25-APR-2000; 2000US-0199384P.  
XX  
PR 21-DEC-2000; 2000US-0256931P.  
XX

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
WI; 2001-616755/71.  
DR N-PSDB; ABA03057.  
XX  
PT Albumin fusion proteins comprising a therapeutic protein and albumin,  
PT useful in the treating immune system disorders (e.g. transplant  
PT rejection), blood related disorders (e.g. myocardial infarction) and  
PT hyperproliferative disorders.  
XX  
PS Claim 1; Fig 15; 606pp; English.  
XX  
CC The present invention relates to albumin fusion proteins, which comprise  
CC a therapeutic protein and albumin. The present sequence is the protein  
CC sequence for mature human serum albumin (HA), which was used to generate  
CC the fusion proteins of the present invention. The albumin fusion proteins  
CC are useful in the treatment, prevention, diagnosis, and/or detection of  
CC diseases/disorders such as immune system disorders (e.g. transplant  
CC rejection), blood related disorders (e.g. myocardial infarction),  
CC hyperproliferative disorders (e.g. childhood acute myeloid leukemia),  
CC renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.  
CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),  
CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders  
CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),  
CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.  
CC irritable bowel syndrome) and wound healing  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 2068; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOQCPEFHVKLNVETFAKTCVADESAE 60  
DB 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOQCPEFHVKLNVETFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTTLEKC 360  
QY 361 CAADPHECYAKVDFDEPKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFDEPKPLVEEPQNL 387  
Search completed: April 19, 2004, 11:51:15  
Job time : 228.626 secs

QY 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEVARRHPDYSVLLRLAKTYETTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEVARRHPDYSVLLRLAKTYETTLEKC 360  
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387

RESULT 13  
AAY84873  
ID AAY84873 standard; protein; 585 AA.  
XX AAY84873;  
DT 08-AUG-2000 (first entry)  
DE Amino acid sequence of a human albumin protein.  
KW Human; albumin; ischemic state; serum protein; metal ion salt;  
KW perioperative ischemia; ischemia; myocardial infarction;  
KW progressive coronary artery disease.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "optionally acetylated, and claimed under claim  
FT 56"

PN WO200020840-A1.  
PD 13-APR-2000.  
XX 01-OCT-1999; 99WO-US022905.  
XX 02-OCT-1998; 98US-00165581.  
PR 02-OCT-1998; 98US-00165926.  
PR 02-OCT-1998; 98US-0102738P.  
PR 11-JAN-1999; 99US-0115392P.  
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.  
PI Bar-Or D, Lau E, Winkler JV;  
XX WPI; 2000-303843/26.

XX New method for the continuous detection of ischemic states comprises  
FT detecting and quantifying the existence of an alteration of the serum  
FT protein albumin.  
XX Disclosure; Page 97-100; 105pp; English.

XX The present sequence represents human albumin protein. The specification  
CC describes a method for the continuous detection of ischemic states. The  
CC method comprises detecting and quantifying the existence of an alteration  
CC of the serum protein albumin. The method comprises contacting a  
CC biological sample containing albumin from the patient with an excess  
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus  
CC of naturally occurring human albumin, to form a mixture containing bound  
CC metal ions and unbound metal ions, and then determining the amount of  
CC metal ions bound to the albumin N-terminus. The amount of bound metal  
CC ions is correlated to a known value to determine the occurrence of non-  
CC occurrence of an ischemic event. The methods are useful for detection of  
CC ischemic states. The methods are also useful for distinguishing  
CC perioperative ischemia from ischemia caused by, amongst other things,  
CC myocardial infarctions and progressive coronary artery disease  
XX Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191; Indels 0; Gaps 0;  
Matches 387; Conservative 0; Mismatches 0;  
QY 1 DAHSEVAHRFKDLGBENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHRFKDLGBENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBERNECFLOHKDDNPNIPLRVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBERNECFLOHKDDNPNIPLRVRPEV 120  
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTJK 240  
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTJK 240  
QY 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEVARRHPDYSVLLRLAKTYETTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEVARRHPDYSVLLRLAKTYETTLEKC 360  
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387

RESULT 14  
AAY83946  
ID AAY83946 standard; protein; 585 AA.  
XX AAY83946;  
DT 28-JUL-2000 (first entry)  
XX Yeast codon-biased recombinant human serum albumin protein.  
DE Yeast codon-biased recombinant human serum albumin protein.  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX Homo sapiens.  
OS Synthetic.  
XX CN1239103-A.  
XX 22-DEC-1999.  
XX 17-JUN-1998; 98CN-00102506.  
XX 17-JUN-1998; 98CN-00102506.  
XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.  
PI Li S, Lu D;  
XX WPI; 2000-351198/31.  
XX N-PSDB; AAA10091.  
XX Process for preparing recombinant human serum albumin comprising yeast  
FT biased sex codons - uses a recombinant DNA technique.  
XX Disclosure; Fig 1; 44pp; Chinese.  
XX The method relates to a method of recombinantly producing human serum  
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as

KW Serum albumin-growth hormone fusion protein; growth hormone;  
KW Down's syndrome.  
XX  
XX Unidentified.  
XX KR99076789-A.  
XX  
XX 15-OCT-1999.  
XX  
XX 25-JUN-1998; 98KR-00704914.  
XX  
XX 30-DEC-1995; 95GB-00026733.  
XX  
XX 19-DEC-1996; 96WO-GB003164.  
XX  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
XX Ballance DJ;  
XX  
XX WPI; 1997-363680/33.  
XX  
XX N-PSDB; AAK95568.  
XX  
XX Serum albumin-growth hormone fusion protein - useful to treat growth  
XX hormone related diseases, e.g. Down's syndrome.  
XX  
XX Disclosure; Fig 6; 21pp; Korean.  
XX  
XX The invention relates to a serum albumin-growth hormone fusion protein -  
XX useful to treat growth hormone related diseases such as Down's syndrome.  
XX This sequence represents a HSA protein related to the serum albumin-  
XX growth hormone protein of the invention  
XX  
XX Sequence 585 AA;  
SQ  
Query Match 100.0%; Score 2068; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCSLQKGFGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKQRLKCSLQKGFGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240  
QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDEMPA 300  
DB 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDEMPA 300  
QY 301 DLPSLAADPVEKQVCKVAEAKDVFGLGFLYFYARRHPDYVSVLLLLRLAKTYETTLK 360  
DB 301 DLPSLAADPVEKQVCKVAEAKDVFGLGFLYFYARRHPDYVSVLLLLRLAKTYETTLK 360  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387  
RESULT 12  
ID AAW59841  
XX AAW59841 standard; protein; 585 AA.  
XX  
XX AAW59841;  
XX AC  
XX  
XX 20-NOV-1998 (first entry)  
DT

XX Mature protein of human serum albumin (HSA).  
DE  
XX  
XX Protein expression; monocotyledon plant cell;  
KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;  
KW ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;  
KW antithrombotic; blood replacement.  
XX  
XX Homo sapiens.  
XX  
XX WO9836085-A1.  
XX  
XX 20-AUG-1998.  
XX  
XX 13-FEB-1998; 98WO-US003068.  
XX  
XX 13-FEB-1997; 97US-0037991P.  
XX  
XX 13-FEB-1997; 97US-0038168P.  
XX  
XX 13-FEB-1997; 97US-0038169P.  
XX  
XX 13-FEB-1997; 97US-0038170P.  
XX  
XX (PHYT-) APPLIED PHYTOLOGICS INC.  
XX  
XX Sutliff TD, Rodriguez RL;  
XX  
XX WPI; 1998-467179/40.  
XX  
XX N-PSDB; AAV41728.  
XX  
XX Expressing mature, glycosylated proteins in monocotyledonous plant cells  
XX - from chimeric gene including signal peptide sequence, specifically  
XX therapeutic agents and industrial enzymes.  
XX  
XX Disclosure; Page 31-32; 53pp; English.  
XX  
XX The present sequence represents the mature protein of human serum albumin  
XX (HSA). The protein is used to exemplify the invention. The specification  
XX describes a method for producing mature heterologous protein in  
XX monocotyledonous plant cells. The method comprises transforming the cells  
XX with a chimeric gene comprising a monocotyledon transcription regulator,  
XX inducible either during seed maturation or by adding/removing a small  
XX molecule, DNA encoding the heterologous protein, and DNA encoding a  
XX signal peptide, with the signal peptide causing secretion of the protein  
XX from the cell. Proteins expressed in this manner include mature  
XX glycosylated alpha 1-antitrypsin (AAT) with a glycosylation pattern that  
XX significantly increases its serum half-life, mature glycosylated  
XX antithrombin III (ATIII), mature human serum albumin (HSA) having the  
XX native folding pattern as shown by bilirubin-binding characteristics, or  
XX mature active subtilisin BPN'. These proteins are useful therapeutically  
XX (e.g. AAT for treating emphysema, ATIII as antithrombotic and HSA as  
XX blood replacement) or as industrial enzymes (BPN' is used in detergents)  
XX  
XX Sequence 585 AA;  
Query Match 100.0%; Score 2068; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCSLQKGFGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKQRLKCSLQKGFGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240



PN JP04211375-A.  
 XX 03-AUG-1992.  
 XX 05-FEB-1991; 91JP-00014600.  
 XX 05-FEB-1990; 90JP-00025682.  
 XX (AJIN ) AJINOMOTO KK.  
 XX WPI; 1992-304940/37.  
 XX N-PSDB; AAQ27813.  
 XX Synthetic gene for prepn. of human serum albumin - comprises synthetic  
 PT DNA contg. gene coding the albumin using coding in Escherichia coli.  
 XX Fig 1; Page 13; 37pp; Japanese.  
 XX The protein sequence was deduced from the synthetic DNA sequence encoding  
 CC human serum albumin which was prepd. by ligating eight synthetic HSA gene  
 CC fragments  
 XX Sequence 585 AA;  
 SQ  
 Query Match 100.0%; Score 2068; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDDNPNLRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDDNPNLRLVRPEV 120  
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
 Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
 QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
 Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
 Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYTTLEKC 360  
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYTTLEKC 360  
 QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
 Db 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
 RESULT 10  
 AAR80301  
 ID AAR80301 standard; protein; 585 AA.  
 XX AAR80301;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 17-JAN-1996 (first entry)  
 XX Human serum albumin.  
 XX Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.  
 XX Homo sapiens.  
 OS  
 XX W09523857-A1.  
 FN

XX 08-SEP-1995.  
 XX 01-MAR-1995; 95WC-QB000434.  
 XX 05-MAR-1994; 94GB-00004270.  
 XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
 XX Kerrywilliams SM, Gilbert SC;  
 XX WPI; 1995-320572/41.  
 XX N-PSDB; AAQ98695.  
 XX Yeast with reduced levels of aspartyl protease 3 proteolytic activity -  
 PT used to secrete human albumin without prodn. of the 45 kD fragment.  
 XX Example 1; Page 26-28; 50pp; English.  
 XX The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected  
 CC to site-directed mutagenesis to investigate the role of endoproteases in  
 CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is  
 CC expressed in S. cerevisiae. Mutations were: R410A; L407A; L408V; V409A;  
 CC and R410A, K413Q, K414Q. The latter set of mutations, especially,  
 CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing  
 CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct  
 CC PI field.)  
 XX Sequence 585 AA;  
 SQ  
 Query Match 100.0%; Score 2068; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDDNPNLRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDDNPNLRLVRPEV 120  
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
 Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
 QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
 Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
 Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYTTLEKC 360  
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYTTLEKC 360  
 QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
 Db 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387  
 RESULT 11  
 AAO20111  
 ID AAO20111 standard; protein; 585 AA.  
 XX AAO20111;  
 AC  
 XX 06-AUG-2002 (first entry)  
 DT  
 XX HSA protein sequence related to the growth hormone protein.  
 DE  
 XX

```
DR N-PSDB; AAQ04719.
XX
XX Human serum albumin prepn. by yeast host - by culturing transformed
PT plasmid yeast to produce serum, and removing it.
XX
XX Disclosure; Page ?; -pp; Japanese.
PS
XX Mature HSA-A may be produced using the sequence incorporated into a
CC plasmid vector with suitable controllers, and transferred to a yeast
CC expression system
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNL 387
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNL 387

RESULT 8
AAR08457
ID AAR08457 standard; protein; 585 AA.
XX
XX AAR08457;
AC
XX
XX 25-MAR-2003 (revised)
DT 16-APR-1991 (first entry)
XX
XX Human serum albumin.
DE
XX HSA; folding; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FT Region 1..303
FT /label= B
FT Region 123..585
FT /label= C
FT Region 123..303
FT /label= A
XX
XX JF02227079-A.
XX
XX 10-SEP-1990.
XX
```

```
PF 25-AUG-1989; 89JP-00217540.
XX
XX 06-OCT-1988; 88JP-00250926.
XX
XX (TOFU ) TONEN CORP.
PA
XX WPI; 1990-317325/42.
DR
XX N-PSDB; AAQ06099.
DR
XX New human serum albumin fragments - used to bond to medicines and for
PT stable folding of protein(s).
XX
XX Claim 1; Fig 8; 24pp; Japanese.
PS
XX Fragments A-C of HSA are expressed as fusion proteins with the signal
CC peptide of E. coli alkaline phosphatase. The fragments are selected for
CC their specific properties. The C-terminal truncated fragment, B, does not
CC bind long-chain fatty acids but does bind to various medicines at the
CC central region. The N-terminal truncated fragment, C, has good stability
CC in protein folding. The central segment, A, has characteristics of both B
CC and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD
CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-
XX 2003 to correct PR field.)
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNL 387
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNL 387

RESULT 9
AAR26362
ID AAR26362 standard; protein; 585 AA.
XX
XX AAR26362;
AC
XX 11-FEB-1993 (first entry)
DT
XX Synthetic HSA protein.
DE
XX Human serum albumin; transformants; recombinant.
XX
XX Synthetic.
OS
XX
```

CC which the codons which are most frequently used by the chosen non-human  
CC host were selected. In this case, it is yeast cells (L120; Leu2-3; 112,  
CC His 3-11, 15). The synthetic HSA gene was assembled from 24  
CC oligonucleotide blocks. HSA is used in therapy for the treatment of  
CC hypovolaemia, shock and hypalbuminaemia. It is also used as an additive  
CC in perfusion liq. for extracorporeal circulation and as an experimental  
CC antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-  
XX -2003 to correct PI field.)  
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240  
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240  
Qy 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Qy 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387

RESULT 6  
AAP90388  
ID AAP90388 standard; protein; 585 AA.  
AC AAP90388;

DT 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 01-NOV-1989 (first entry)  
DE Mature human serum albumin polypeptide.  
DE Human serum albumin; mature protein; new polypeptides; plasma expanders.  
KW Homo sapiens; (Human).  
OS Homo sapiens.  
PN EP322094-A.  
PD 28-JUN-1989.  
PF 25-OCT-1988; 88EP-00310000.  
PR 30-OCT-1987; 87GB-00025529.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PI Ballance DJ, Hinchliffe E, Gelsow MJ, Senior PJ;  
XX WPI; 1989-186464/26.

DR N-PSDB; AAN90128.  
XX New N-terminal fragments of human serum albumin - esp. useful as blood  
PT plasma expanders.  
XX Disclosure; Fig 2; 20pp; English.

CC Mature protein of human serum albumin (see corresp. AAN90128). Used to  
CC make new N-terminal fragments which are used as plasma expanders, or as  
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-WAR-  
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS  
XX field)  
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.1e-191;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240  
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240  
Qy 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Qy 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387

RESULT 7  
AAR05318  
ID AAR05318 standard; protein; 585 AA.  
AC AAR05318;

DT 08-OCT-1990 (first entry)  
DE Human serum albumin gene product.  
DE Human serum albumin; HSA-A; yeast; ds.  
KW Homo sapiens.  
OS Homo sapiens.  
PN JP02117384-A.  
PD 01-MAY-1990.  
PF 26-OCT-1988; 88JP-00268302.  
PR 26-OCT-1988; 88JP-00268302.  
PA (TOFU ) TOA NENRYO KOGYO KK.  
XX WPI; 1990-176228/23.

| Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |   |
|--|---|
| Qy   | 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPEEDHVKLVNEVTEFAKTCVADESA 60         |
| Db   | 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPEEDHVKLVNEVTEFAKTCVADESA 60         |
| Qy   | 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPEV 120     |
| Db   | 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPEV 120     |
| Qy   | 121 DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPPELLFPAKRYKAAPTECCQADKAACLLP 180 |
| Db   | 121 DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPPELLFPAKRYKAAPTECCQADKAACLLP 180 |
| Qy   | 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLT 240     |
| Db   | 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLT 240     |
| Qy   | 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSKHSCIAEVENDEMPA 300   |
| Db   | 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSKHSCIAEVENDEMPA 300   |
| Qy   | 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEK 360   |
| Db   | 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEK 360   |
| Qy   | 361 CAADPHECYAKVDFEFPKLVPEPQNL 387                                    |
| Db   | 361 CAADPHECYAKVDFEFPKLVPEPQNL 387                                    |

RESULT 4  
AAP90392 ID AAP90392 standard; protein; 407 AA.  
XX AC AAP90392;  
XX DT 24-OCT-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 01-NOV-1989 (first entry)  
XX DE N-terminal human serum albumin.  
XX DE Human serum albumin; mature protein; new polypeptides; plasma expanders.  
XX OS Homo sapiens; (Human).  
XX PN EP322094-A.  
XX PD 28-JUN-1989.  
XX PF 25-OCT-1988; 88EP-00310000.  
XX PR 30-OCT-1987; 87GB-00025529.  
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
XX DR WPI; 1989-186464/26.  
XX PT New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.  
XX PS Claim 2; Page 9; 20pp; English.  
XX CC N-terminal portion of human serum albumin. Used to make plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25 -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)  
XX SQ Sequence 407 AA;

Query Match 100.0%; Score 2068; DB 1; Length 407;  
Best Local Similarity 100.0%; Pred. No. 6.9e-192;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPEEDHVKLVNEVTEFAKTCVADESA 60  
Db 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPEEDHVKLVNEVTEFAKTCVADESA 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRVPEV 120  
Qy 121 DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPPELLFPAKRYKAAPTECCQADKAACLLP 180  
Db 121 DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPPELLFPAKRYKAAPTECCQADKAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSKHSCIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSKHSCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEK 360  
Qy 361 CAADPHECYAKVDFEFPKLVPEPQNL 387  
Db 361 CAADPHECYAKVDFEFPKLVPEPQNL 387  
RESULT 5  
AAP93344 ID AAP93344 standard; protein; 585 AA.  
XX AC AAP93344;  
XX DT 25-MAR-2003 (revised)  
XX DT 23-JUN-1990 (first entry)  
XX DE Sequence of mature human serum albumin (HSA) as encoded by artificial Gene.  
XX DE Mature human serum albumin; artificial gene; oligonucleotide block; hypobolaemia; shock; hypobolaemia.  
XX OS Homo sapiens.  
XX PN EP308381-A.  
XX PD 22-MAR-1989.  
XX PF 13-SEP-1988; 88EP-00850299.  
XX PR 14-SEP-1987; 87SE-00003539.  
XX PA (SKAN-) SKANDIGEN AB.  
XX PA (MAGY ) MTA SZEGEDI BIOLOG KOEZFONTI.  
XX PA (VEPE-) VEPEX CONTRACTOR LTD.  
XX PI Aberg B, Simoncsits A, Kalan M, Csperpan I, Bajszar G;  
XX DR WPI; 1989-087749/12.  
XX DR N-PSDB; AAN90997.  
XX PT Artificial gene coding for authentic human serum albumin - constructed on the basis of codons most frequently used by chosen non-human host.  
XX PS Disclosure; pp. 11-16; 121pp; English.  
XX CC The synthetic gene was constructed by designing a nucleotide sequence in

Db 1 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTTECCQADKAAACLLP 180  
Db 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTTECCQADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFFPKAEFAEVS KLVTDLTK 240  
Db 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFFPKAEFAEVS KLVTDLTK 240  
QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 2  
AAP90390  
ID AAP90390 standard; protein; 389 AA.  
XX AC AAP90390;  
XX DT 24-OCT-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 01-NOV-1989 (first entry)  
XX DE N-terminal human serum albumin.  
XX KW N-terminal portion of human serum albumin; plasma expanders.  
XX OS Homo sapiens; (Human).  
XX PN EP322094-A.  
XX PD 28-JUN-1989.  
XX PF 25-OCT-1988; 88EP-00310000.  
XX PR 30-OCT-1987; 87GB-00025529.  
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
XX DR WPI; 1989-186464/26.  
XX PT New N-terminal fragments of human serum albumin - esp. useful as blood  
XX PS plasma expanders.  
XX PS Claim 2; Page 9; 20pp; English.  
XX CC N-terminal portion of human serum albumin. Used to make plasma expanders,  
XX CC or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25  
XX CC -MAR-2003 to correct FA field.) (Updated on 24-OCT-2003 to standardise OS  
XX CC field)  
XX SQ Sequence 389 AA;  
Query Match 100.0%; Score 2068; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 6.4e-192;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
Db 1 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTTECCQADKAAACLLP 180  
Db 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTTECCQADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFFPKAEFAEVS KLVTDLTK 240  
Db 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFFPKAEFAEVS KLVTDLTK 240  
QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 3  
AAP90391  
ID AAP90391 standard; protein; 390 AA.  
XX AC AAP90391;  
XX DT 24-OCT-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 01-NOV-1989 (first entry)  
XX DE N-terminal human serum albumin.  
XX KW N-terminal portion of human serum albumin; plasma expanders.  
XX OS Homo sapiens; (Human).  
XX PN EP322094-A.  
XX PD 28-JUN-1989.  
XX PF 25-OCT-1988; 88EP-00310000.  
XX PR 30-OCT-1987; 87GB-00025529.  
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
XX DR WPI; 1989-186464/26.  
XX PT New N-terminal fragments of human serum albumin - esp. useful as blood  
XX PS plasma expanders.  
XX PS Claim 2; Page 9; 20pp; English.  
XX CC N-terminal portion of human serum albumin. Used to make new N-terminal  
XX CC fragments which are used as plasma expanders, or as substitutes for HSA  
XX CC or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct FA  
XX CC field.) (Updated on 24-OCT-2003 to standardise OS field)  
XX SQ Sequence 390 AA;  
Query Match 100.0%; Score 2068; DB 1; Length 390;  
Best Local Similarity 100.0%; Pred. No. 6.5e-192;



|    |     |   |     |
|----|-----|---|-----|
| Db | 215 | LNEHVCSVIRKFSRNLQATTIIKLSQKLTAEANFTIQLALDVALIHERCCQNSLECL   | 274 |
| QY | 255 | DDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPADLPSLAADFVESKD | 314 |
| Db | 275 | QDGEKWTYICQQNLTSSKIAECCCKPMIQLGFCIIHAENGVKPEGLSLNPSQFLGDRN  | 334 |
| QY | 315 | VCKNYAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETTLKCCCAADPHECYAKVF   | 374 |
| Db | 335 | FAQFSSEKIMFMASFHEYSETHENLPVSVILRIAKTYQETLEKCSQSGNLPQCQDNLE  | 394 |
| QY | 375 | DEFKPLVEBPQNL   | 387 |
| Db | 395 | BELOKHIESQAL  | 407 |

Search completed: April 19, 2004, 11:59:57  
Job time : 151.368 secs

```

Db      87 DMEGLPQKH-NFSSHCCGKAGPRRLCFFYNKANKVGFLLPPTL-----DPEKCKQAKYKNS 142
QY      132 ETLKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLPKXDELRLDEGKA 191
Db      143 ESFLHLYMYEVARPNFVFAPVLLAAVAWFEEAATTCCEQQOKATCFQAKAAPITQYLKA 202
QY      192 SSQKRLKCASTOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDL 251
Db      203 SSSYQRNVCGALIKFGPKVLNINVAVFSKFPKIGFKDLTLLSDVSSMYEGCGEDV 262
QY      252 ECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPISLAADFVE 311
Db      263 HCIRSOQVNHICSKQDSISSKIKVCEKTLEREACIINANKDDRPEGLSLREAKFTE 322
QY      312 SKDVCNRYAEADVFLGFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db      323 SENVCQERDSDPKFPAEFAEFAEVSGLVTLTKVHTECHGDL 382
QY      372 KVFDEF 377
Db      383 HVEDKF 388

```

RESULT 14

```

Q8UW06 PRELIMINARY; PRT; 624 AA.
AC Q8UW06;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Serum albumin precursor.
OS ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR GO; GO:0003615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

```

```

Query Match 30.5%; Score 630; DB 13; Length 624;
Best Local Similarity 33.7%; Pred. No. 4.1e-46;
Matches 125; Conservative 66; Mismatches 170; Indels 10; Gaps 3;

QY 14 LGENFKALVLIJAFAYLQCCPFEDHVKLVNVEVFATCTVADESAENCDSLHTLFGDK 73
Db 45 IGVHAKALAMALFQMLSKCPHHEGVQVRVNMADI LCSGAKHGDCGKSVITILNE 104
QY 74 LCTVATLRETYGEMADCCAKQBPNERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEE 132
Db 105 ICTPENPEKYPHPEGCKCKEDPERHKCFIEHKSITDPKERTYVVKPSPEQICKDAENRD 164
QY 133 TELKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLPKDELRLDEGKA 191

```

```

Db      165 EPLGHVTHKVASSTHTMYPPAILSLFTLHFDGIVSHCCCKDEATVGQCLSEKMPAKHEVEH 224
QY      192 SSQKRLKCASTOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDL 251
Db      225 VCAVQKHNCVILQNFNERALRASKAHAHACSKFPHASFENVRQLTDGIVHLHQTCCGDDMM 284
QY      252 ECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPISLAADFVE 311
Db      285 ACVAERKMLTQTCEK-----KCCCKPVLSECEIVLPENDEKPADLSEVRYFDP 336
QY      312 SKDVCNRYAEADVFLGFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db      337 DPEVCRFKEEGDAFMRFLCDYAKIHPBSAELNRIASGLEKAYKTCGAGEAHNECIA 396
QY      372 KVFDEFKPLVE 382
Db      397 KEEETLRHEIE 407

```

RESULT 15

```

Q8BK65 PRELIMINARY; PRT; 605 AA.
AC Q8BK65;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Alpha-fetoprotein precursor.
OS AFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK076053; BAC36150.1; -.
DR MGI; MGI:87951; Afp.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0001542; P:ovulation (sensu Mammalia); IMP.
DR GO; GO:0042448; P:progesterone metabolism; IMP.
DR GO; GO:0019953; P:sexual reproduction; IMP.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 605 AA; 67322 MW; 048B7A4A8B01EA4B CRC64;

```

```

Query Match 29.7%; Score 615; DB 11; Length 605;
Best Local Similarity 32.4%; Pred. No. 7.9e-45;
Matches 121; Conservative 80; Mismatches 166; Indels 6; Gaps 3;

QY 16 EENFKALVLIJAFAYLQCCPFEDHVKLVNVEVFATCTVADESAENCDSLHTLFGDKLC 75
Db 40 EKNVLIATITFTQFVPEATEEE---VNMKTSVDLAAKMKNSGDCLESQLSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQBPNERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEET 134
Db 96 HETELSKYQ-LGSCCSQSGVERHQCLLAARKKTAPASVFPFQPEAESCAHEERAVTF 154
QY 135 LKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLPKDELRLDEGKASSA 194
Db 155 MNRFIYVSRRNPFMYAPAILSLAAQYDKVVLACKADNKECECFQTKRASIAKELREGSM 214
QY 195 KQRLKCASTOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDLLECA 254

```



```

RP SEQUENCE FROM N.A.
RA Kim J.G., Noneman D., Vallet J.L., Christenson R.K.:
RT "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSCs.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517770; AAM66710.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; P:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0006264; Serum albumin.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match 33.7%; Score 697.5; DB 6; Length 610;
Best Local Similarity 36.4%; Pred. No. 5.5e-52;
Matches 136; Conservative 74; Mismatches 161; Indels 3; Gaps 3;

QY 16 EENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESSENCDKSLHTLFGDKLC 75
DB 40 ENMLVDLATIFFAQFVHGATYTEIRKQANDVLAVIKPTDGHPTGCLENQMSAFLEBIC 99
QY 76 TVATLRETYGEMADCCAKOEPERNECFLOHKDDNP-NLPLRVPEVDVMTAFHDNEETP 134
DB 100 HSEIEPEKYG-LSHCCSQSGEERHNCFLARKKAAPASIPFFQVPEFVTSKAYEENRELF 158
QY 135 LKLYEIAARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLPKLDELDRDEGKASSA 194
DB 159 MTEYIIEIARRHPFLVAPILSLAAGYDKIIPCCKAENAVECFQTKAASITKELRESSL 218
QY 195 KQRLKASLOKGERAFKAWAVARLSORFPKAEFVSKLVTLTKVHTCCCHGDLLECA 254
DB 219 LNCHMCTVMRQFGARTFRAITVTKLSQKPKANFTEIQKLVLDVAHIHECCRGNVLESL 278
QY 255 DDRADLAKYCENODSISKKLCECKP-LLEKSHCIAEVDENDEMPADLPISLAADFVSK 313
DB 279 QDAERVVSVCSQDITLSSKIAECCKLTLELGQCIHAENDDKPEGLSPNLNFI GER 338
QY 314 DVCKNVAEKDVFLGMFLYEAARRHPDYVSVLLLRLLAKTYETTLKCCAAADPHCYAKV 373
DB 339 DFNQLSREKDLGMAFTVEYSRRHPKLAVPVILRVKGYQELLEKCSQSENPLECDKG 398
QY 374 FDEFKPLVEBPQNL 387
DB 399 EEELEKYIQESQAL 412

RESULT 12
Q7TSF3 PRELIMINARY; PRT; 608 AA.
AC Q7TSF3
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Alpha-fetoprotein.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
[1]
SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Lin H.L., Chen P.J., Wu H.L., Lee R.S., Chen D.S.:
RT "Molecular Cloning and Expression of Woodchuck Alpha-Fetoprotein.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034049; AAK55757.1; -.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 67962 MW; EB46D7A37382B222 CRC64;

Query Match 33.2%; Score 686.5; DB 11; Length 608;
Best Local Similarity 36.8%; Pred. No. 5e-51;
Matches 137; Conservative 64; Mismatches 157; Indels 8; Gaps 3;

QY 16 EENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESSENCDKSLHTLFGDKLC 75
DB 27 DENTAYLAIASFQVQESAFDEVELVKWMLDYDRCDWADNTLPECSKTANDAIQDWLC 86
QY 76 TVATLRETYGEMADCCAKOEPERNECFLOHKDDN-----PNLPLRVPEVDVMTAFHDNE 131

```

```

QY 116 VRPEVDVMTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCQCAADKA 175
DB 148 VRPAEQICCKDAENRGIPILARYIFMLALGHPHYIPAILGFAQFQDGIIVSHCKDVETA 207
QY 176 A-CLLPKLDLRLDEKASAKQRLKASLOKFGGERAFKAWAVARLSQRFPKAEFAYSKL 234
DB 208 GQCFNDKMPKHEQVEYVYCALQKHCNYILOQFKERALYAKAVQASQKFFPLASFENQII 267
QY 235 VTDLTKVTECHGDLLECCADRADLAKVICENODSISKKKECEKPLLEKSHCIAEVE 294
DB 268 VPDTHLHQTCGGGMACMLERMKLTAKICEKDELAIHLKECCDKELLERSACIIRLP 327
QY 295 NDEMPADLPSLAADFVESKOVCKNYAEAKDVLGMLFYEARHPDYVSVILLRLAKTYE 354
DB 328 NDKQADLPSPKPVHYIDDPEVCKLYTEGGDTFMGRFLYECARRHODYSPEMLLAWGSGYE 387
QY 355 TTEKCCAAADPHECYAKVDFEFKPLVEPQNL 387
DB 388 EFLKCCCAAGHNECLAETEESLKEIBSVTL 420

RESULT 9
Q8CG74 PRELIMINARY; PRT; 205 AA.
AC Q8CG74;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvetTacrB;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A2277794; CAC81903.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 1.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
FT NON_TER 205
SQ SEQUENCE 205 AA; 23624 MW; 0A982B85C7474BF9 CRC64;

Query Match 35.7%; Score 739; DB 11; Length 205;
Best Local Similarity 69.6%; Pred. No. 3.7e-56;
Matches 126; Conservative 26; Mismatches 29; Indels 0; Gaps 0;

QY 1 DAHSEVNRFDLGEENFKALVLIATAQYLOQCPFDHVKLNVETFEAKTCVADESAG 60
DB 25 EAHKSEIARYNDLGEQHFKGLVLIATFSAQYLOKCSYDEHAKLVQEVTDPTFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDNPMLPRLVREPV 120
DB 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECFQHKDNPSPFPERPEA 144
QY 121 DVNCTAFHNDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCQCAADKAACLLP 180
DB 145 EAMCTSFKENPTTFMGHLYHEVARRHPYFYAPPELLYAEQYNEILTQCCAEADKESCLTP 204
QY 181 K 181
DB 205 K 205

RESULT 10
Q8MJU5 PRELIMINARY; PRT; 609 AA.
ID Q8MJU5

```

```

AC Q8MJU5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Alpha-fetoprotein.
GN AFP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furuichi M., Neo S., Hisasue M., Teuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.;
RT "Canine alpha-fetoprotein cDNA."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 3.
DR ProDom; PD002486; Serum albumin; 1.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250CSAF2AF0 CRC64;

Query Match 34.4%; Score 712; DB 6; Length 609;
Best Local Similarity 37.3%; Pred. No. 3.1e-53;
Matches 139; Conservative 70; Mismatches 162; Indels 2; Gaps 2;

QY 16 BENFKALVLIATAQYLOQCPFDHVKLNVETFEAKTCVADESAGNCDKSLHTLFGDKLC 75
DB 40 EMNLVDLATIFFAQFVQEAITYKEVSKVKDILTVIERKSTGSEQFGGLENQLAFLEEC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECEFLQHKD-DNPNLPRLVREPVDMCTAFHNDNEET 134
DB 100 HEKEISEKYG-LADCCSQREERHNCFLAHKKAAPSPFPQVAEPVTSCKAYEENRDMF 158
QY 135 LKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCQCAADKAACLLPKLDLRLDEGKASSA 194
DB 159 MNRYIYEIARRHPFLYAPTILSLAAHYGKIIP-CCKAENAVECFQTKTSLITKELRESSL 218
QY 195 KQRLKCAKSLQFGERAFKAWAVARLSORFPKAEFAVSVLTLTKVHTSCCHGDLLECA 254
DB 219 LNQHICAVMENFGPRTFRAITVTKLSQKFSKANFTIQKLVLVDVAIHBECCRGVLECL 278
QY 255 DADRADLAKYICENODSISKKKECEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
DB 279 QDGEKIMSYICSQODILSSKIADCCCKLPILLEGQCIHAENDGKPEGLSPNLNRFLEED 338
QY 315 VCKNYAEAKDVLGMLFYEARHPDYVSVLLRLAKYVETTTLEKCCAAADPHECYAKVF 374
DB 339 FNQSSREKDLFWARFYEYRRRHTKLVAPVVLVAVKGYOELLEKKSQSENPLECQDKGE 398
QY 375 DEFKPLVEEPQNL 387
DB 399 EELEKYIQESQAL 411

RESULT 11
Q8MJU76 PRELIMINARY; PRT; 610 AA.
AC Q8MJU76;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;

```

|  |   |     |   |     |
|--|---|-----|---|-----|
| DR   | Pfam; PF00273; transport_prot; 2.                                 | 62  | CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPN-LPRLVRPEV | 120 |
| DR   | PRINTS; PR00802; SERUMALBUMIN.                                    |     |   |     |
| DR   | SMART; SM00103; ALBUMIN; 2.                                       | 5   | CLASLDTIFLDEICHHEGFAAKY-DLAACCAAEVERKECLLAHKUATGFIQAFORPGI    | 63  |
| DR   | PROSITE; PS00212; ALBUMIN; 2.                                     |     |   |     |
| SQ   | SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;                |     |   |     |
| Query Match 45.3%; Score 936; DB 4; Length 396;                  |   |     |   |     |
| Best Local Similarity 49.9%; Pred. No. 6.7e-73;                  |   |     |   |     |
| Matches 196; Conservative 42; Mismatches 121; Indels 34; Gaps 7; |   |     |   |     |
| QY   | 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE    | 60  |   |     |
| DB   | 25 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE   | 84  |   |     |
| QY   | 61 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPNLPRLVRPEV  | 120 |   |     |
| DB   | 85 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPNLPRLVRPEV  | 144 |   |     |
| QY   | 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP | 180 |   |     |
| DB   | 145 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP | 183 |   |     |
| QY   | 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSXKLVTDL    | 238 |   |     |
| DB   | 184 VFDEFKPLVEEPQNLIKQ--NCLFQGLGEYFQNALVRYTKVPQVSTPTLVEVSRL       | 241 |   |     |
| QY   | 239 TKVHTBCC---HGDILLECADDRADLA-KYIC--ENQDSISSKLKCECKEPLLEKSHCIAE | 292 |   |     |
| DB   | 242 GKVGSKCKCKHPEAKRMPCEADYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPQPSA  | 301 |   |     |
| QY   | 293 VENDEMPADPLSLAADFVESKVKYAKVFLGMFLYIYARRHPDYSVVLRLAKT          | 352 |   |     |
| DB   | 302 LEVDETYVYKFEKNAFTTFHADICTLSERQIKKQTALVELVHKPKATKEQLKAVMD      | 361 |   |     |
| QY   | 353 YETTLKCCAAADPHECYAKVDFEKFPLVEEPQ 385                          |     |   |     |
| DB   | 362 FAAFVEKCKCKADKTCFA---EEGKLVAASQ 391                           |     |   |     |
| RESULT 7   |   |     |   |     |
| ID   | Q8JIA9 PRELIMINARY; PRT; 527 AA.                                  |     |   |     |
| AC   | Q8JIA9;   |     |   |     |
| DT   | 01-OCT-2002 (TrEMBLrel. 22, Created)                              |     |   |     |
| DT   | 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)                 |     |   |     |
| DT   | 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)               |     |   |     |
| DE   | Serum albumin (Fragment).   |     |   |     |
| OS   | Sphenodon punctatus (Hatteria) (Tuatara).                         |     |   |     |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |     |   |     |
| OC   | Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.            |     |   |     |
| OX   | NCBI_TaxID=8508;  |     |   |     |
| RN   | [1]   |     |   |     |
| RP   | SEQUENCE FROM N.A.  |     |   |     |
| RA   | Metcalif V.J., Brennan S.O., George P.M., Chambers G.K.;          |     |   |     |
| RA   | "Partial coding sequence for Sphenodon punctatus 68 kDa albumin." |     |   |     |
| RL   | Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.           |     |   |     |
| DR   | EMBL; AF375971; AAM46104.1; -                                     |     |   |     |
| DR   | GO; GO:0005615; C:extracellular space; IEA.                       |     |   |     |
| DR   | GO; GO:0005386; P:carrier activity; IEA.                          |     |   |     |
| DR   | GO; GO:0006810; P:transport; IEA.                                 |     |   |     |
| DR   | InterPro; IPR00264; Serum_albumin.                                |     |   |     |
| DR   | PRINTS; PR00802; SERUMALBUMIN.                                    |     |   |     |
| DR   | ProDom; PD002486; Serum_albumin; 1.                               |     |   |     |
| DR   | SMART; SM00103; ALBUMIN; 3.                                       |     |   |     |
| DR   | PROSITE; PS00212; ALBUMIN; 3.                                     |     |   |     |
| FT   | NON_TER 1   |     |   |     |
| SQ   | SEQUENCE 527 AA; 59711 MW; C62B799B3875929 CRC64;                 |     |   |     |
| Query Match 38.2%; Score 791; DB 13; Length 527;                 |   |     |   |     |
| Best Local Similarity 45.3%; Pred. No. 3.6e-60;                  |   |     |   |     |
| Matches 148; Conservative 53; Mismatches 124; Indels 2; Gaps 2;  |   |     |   |     |
| QY   | 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE    | 60  |   |     |
| DB   | 25 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE   | 84  |   |     |
| QY   | 61 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPNLPRLVRPEV  | 120 |   |     |
| DB   | 85 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPNLPRLVRPEV  | 144 |   |     |
| QY   | 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP | 180 |   |     |
| DB   | 145 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP | 183 |   |     |
| QY   | 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSXKLVTDL    | 238 |   |     |
| DB   | 184 VFDEFKPLVEEPQNLIKQ--NCLFQGLGEYFQNALVRYTKVPQVSTPTLVEVSRL       | 241 |   |     |
| QY   | 239 TKVHTBCC---HGDILLECADDRADLA-KYIC--ENQDSISSKLKCECKEPLLEKSHCIAE | 292 |   |     |
| DB   | 242 GKVGSKCKCKHPEAKRMPCEADYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPQPSA  | 301 |   |     |
| QY   | 293 VENDEMPADPLSLAADFVESKVKYAKVFLGMFLYIYARRHPDYSVVLRLAKT          | 352 |   |     |
| DB   | 302 LEVDETYVYKFEKNAFTTFHADICTLSERQIKKQTALVELVHKPKATKEQLKAVMD      | 361 |   |     |
| QY   | 353 YETTLKCCAAADPHECYAKVDFEKFPLVEEPQ 385                          |     |   |     |
| DB   | 362 FAAFVEKCKCKADKTCFA---EEGKLVAASQ 391                           |     |   |     |
| RESULT 8   |   |     |   |     |
| ID   | Q8UW05 PRELIMINARY; PRT; 626 AA.                                  |     |   |     |
| AC   | Q8UW05;   |     |   |     |
| DT   | 01-MAR-2002 (TrEMBLrel. 20, Created)                              |     |   |     |
| DT   | 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)                 |     |   |     |
| DT   | 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)               |     |   |     |
| DE   | Serum albumin precursor.  |     |   |     |
| GN   | ALB.  |     |   |     |
| OS   | Ambystoma maculatum (Spotted salamander).                         |     |   |     |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |     |   |     |
| OC   | Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;     |     |   |     |
| OC   | Ambystoma.  |     |   |     |
| OX   | NCBI_TaxID=43114;   |     |   |     |
| RN   | [1]   |     |   |     |
| RP   | SEQUENCE FROM N.A.  |     |   |     |
| RA   | Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;         |     |   |     |
| RA   | "Serum albumin of the mole salamanders Ambystoma maculatum and    |     |   |     |
| RL   | Ambystoma texanum."   |     |   |     |
| RL   | Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.           |     |   |     |
| DR   | EMBL; AF217183; AAL56646.1; -                                     |     |   |     |
| DR   | GO; GO:0005615; C:extracellular space; IEA.                       |     |   |     |
| DR   | GO; GO:0005386; P:carrier activity; IEA.                          |     |   |     |
| DR   | GO; GO:0006810; P:transport; IEA.                                 |     |   |     |
| DR   | InterPro; IPR00264; Serum_albumin.                                |     |   |     |
| DR   | Pfam; PF00273; transport_prot; 3.                                 |     |   |     |
| DR   | PRINTS; PR00802; SERUMALBUMIN.                                    |     |   |     |
| DR   | ProDom; PD002486; Serum_albumin; 1.                               |     |   |     |
| DR   | SMART; SM00103; ALBUMIN; 3.                                       |     |   |     |
| DR   | PROSITE; PS00212; ALBUMIN; 2.                                     |     |   |     |
| FT   | SIGNAL 1 24 POTENTIAL.  |     |   |     |
| FT   | CHAIN 25 626 SERUM ALBUMIN.                                       |     |   |     |
| FT   | CHAIN 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;                   |     |   |     |
| SQ   | SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;                |     |   |     |
| Query Match 36.9%; Score 764; DB 13; Length 626;                 |   |     |   |     |
| Best Local Similarity 38.9%; Pred. No. 9.7e-58;                  |   |     |   |     |
| Matches 153; Conservative 64; Mismatches 170; Indels 6; Gaps 3;  |   |     |   |     |
| QY   | 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE    | 56  |   |     |
| DB   | 28 EGHVDNPPHLIGDLIPMIGVDNSKGLVLAASQMLPLCPYEHLQVRVDVNAQIADLCAGK    | 87  |   |     |
| QY   | 57 ESEAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEKPERNECFLOHKDDNPNLPRL  | 115 |   |     |
| DB   | 88 ARHANCAPSPWTLIDELCKKPEKNAEYFFHQECCCKEDPERHKCFVEHKANHEELIKY     | 147 |   |     |

```
QY 361 CAADPHCYAKVDFEKPVEEPQNL 387
DB 385 CEAENPRACYGVLAFAEPQVEEPQNL 411

RESULT 4
Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR NON_TER 1
FT SEQUENCE 576 AA; 65002 MW; F8573B99AE37F04 CRC64;
SQ

Query Match 74.6%; Score 1542; DB 11; Length 576;
Best Local Similarity 71.5%; Pred. No. 2.8e-125;
Matches 271; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 9 HRFKDLGEENKALVIAFAQVLCQCPEDHVKLVNEVTEFAKTCVADESACNCKSLHT 68
DB 1 NRINDLGEQHFKGLVLIAFSQVLCQCSYDEHAKLVQEVTDFAKTCVADESACNCKSLHT 60

QY 69 LFGDKLCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLPRLVPEVDMCTAFH 128
DB 61 LFGDKLCAIPNLRENYGELADCTCKOEPERNECFQHKDDNPNLPPEPPEAEMCTGFK 120

QY 129 DNEETPLKYLVEIARRHRYFYAPELLFFAKRYKAAPTECCOAKAACLPLKDLRDE 188
DB 121 ENPTTFMGHYLVEHARRHYFYAPELLFYAQNELITOCCEADKESCLTPLDGVKKEK 180

QY 189 GRASSAKQRLKASLQKFGERAFKAWARLSQRFPAEFAVSQVLTDLTKVHTECCGG 248
DB 181 ALVSSVRQRMKSSMQKFGERAFKAWARLSQTFPNAADFARITKLTATDLTKVKECCGG 240

QY 249 DLLECAADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPADLPSLAAD 308
DB 241 DLLECAADDRADLAKYICENQATISSKLTQCCDKPLLKAKHCLSEVEHDTMPADLPAIAD 300

QY 309 FVSKDVKCNKAAKDVFLGMFLYEVARRHPDYSVVLRLAKYVETTTLEKCAAADPHE 368
DB 301 FVEDQEVCKNKAFAKDVFLGTFLYEVARRHPDYSVSLRLAKYATLEKCAANPPA 360

QY 369 CYAKVDFEKPVEEPQNL 387
DB 361 CYGTVLAEPQVEEPQNL 379
```

```
RESULT 5
Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
DR SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;
SQ

Query Match 46.2%; Score 956; DB 4; Length 417;
Best Local Similarity 90.6%; Pred. No. 1.3e-74;
Matches 183; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

QY 186 RDEGKASSAKQRLKASLQKFGERAFKAWARLSQRFPAEFAVSQVLTDLTKVHTEC 245
DB 24 RDAHSEVAHR-----FKDLGEENKAWARLSQRFPAEFAVSQVLTDLTKVHTEC 77

QY 246 CHGDLLECAADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPADLPSL 305
DB 78 CHGDLLECAADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPADLPSL 137

QY 306 AADFVESKDVCKNKAFAKDVFLGMFLYEVARRHPDYSVVLRLAKYVETTTLEKCAAAD 365
DB 138 AADFVESKDVCKNKAFAKDVFLGMFLYEVARRHPDYSVVLRLAKYVETTTLEKCAAAD 197

QY 366 PHECYAKVDFEKPVEEPQNL 387
DB 198 PHECYAKVDFEKPVEEPQNL 219

RESULT 6
Q81UK7 PRELIMINARY; PRT; 396 AA.
AC Q81UK7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR SEQUENCE 396 AA; 47360 MW; 16E764833EEF4E8D CRC64;
SQ
```

```

Db 121 DAMCTAFHENEQFLGKLYEIARRHPYFYAPPELLYAAEYRGVFTCECAADKAACLTLP 180
Qy 181 KLDELDRDEGKASSAKORLKASLQKFGERAFAKAWAVARLSQRFKPAFAEVSKLVTDLTK 240
Db 181 KYDALREKVLASSAKERLKASLQKFGERAFAKAWAVARLSQRFKPAFAEVSKLVTDLTK 240
Qy 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 241 IHKECCHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVSKDYCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 360
Db 301 DLPLLAADFVEDKEVKYQAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 360
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CATDPPACAYAHVDFEFKPLVEEPHNL 387

```

## RESULT 2

```

Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staderker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0003386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; 55EAB28E1C66354 CRC64;

```

```

Query Match 81.3%; Score 1681; DB 5; Length 608;
Best Local Similarity 76.5%; Pred. No. 2.6e-137;
Matches 296; Conservative 53; Mismatches 38; Indels 0; Gaps 0;

```

```

Qy 1 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQCCPFEDHVKLNVETFAKTCVADESAAE 60
Db 25 DAHSEIAHRYNDLGEQHFGLVLIAPSOFLQCPYEHEHVKLNVETDFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCAIPLRLDSYGLADCCAKQEPERNECFLOHKDDNPNLPPFVRPDA 144
Qy 121 DVNCTAFHDNEETFLKYLIEIARRHPYFYAPPELLYFAKRYKAFTCCQAADKAACLTLP 180
Db 145 EAMCTSFQENAVTFMGHYLHEVARHPYFYAPPELLYAEKYSAIMTECCGEADKAACITP 204
Qy 181 KLDELREGKASSAKORLKASLQKFGERAFAKAWAVARLSQRFKPAFAEVSKLVTDLTK 240
Db 205 KLDALKEKALASSVNRQLKCSLQRFQORAFKAWAVARNSQKPKADFAEITKLATDLTK 264
Qy 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 LTERCCHGDLLECADRADLAKYICENQDSISSKLQACCDKPVLLKSHCLSEVENDDLPA 324

```

```

Qy 301 DLPSLAADFVSKDYCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 360
Db 325 DLPSLAADFVEDKEVKYQAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 384
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 385 CABADSPSACYGKVLDFEFQPLVEEPKNL 411

```

## RESULT 3

```

Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3
AC Q8C7H3
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RC MEDLINE=23354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0003386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

```

```

Query Match 76.6%; Score 1584; DB 11; Length 608;
Best Local Similarity 71.8%; Pred. No. 6.8e-129;
Matches 278; Conservative 56; Mismatches 53; Indels 0; Gaps 0;

```

```

Qy 1 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQCCPFEDHVKLNVETFAKTCVADESAAE 60
Db 25 DAHSEIAHRYNDLGEQHFGLVLIAPSOFLQCPYEHEHVKLNVETDFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCAIPLRLDSYGLADCCAKQEPERNECFLOHKDDNPNLPPFVRPDA 144
Qy 121 DVNCTAFHDNEETFLKYLIEIARRHPYFYAPPELLYFAKRYKAFTCCQAADKAACLTLP 180
Db 145 EAMCTSFQENAVTFMGHYLHEVARHPYFYAPPELLYAEKYSAIMTECCGEADKAACITP 204
Qy 181 KLDELREGKASSAKORLKASLQKFGERAFAKAWAVARLSQRFKPAFAEVSKLVTDLTK 240
Db 205 KLDGVEKALVSSVNRQKSSMQKFGERAFAKAWAVARLSQTFNADFAITKLATDLTK 264
Qy 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VNKECCHGDLLECADRADLAKYICENQATISSKLQCTCCDKPLLKAKHCLSEVEDHTMPA 324
Qy 301 DLPSLAADFVSKDYCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 360
Db 325 DLPLAADFVEDQEVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKYETTLTKC 384

```



|    |     |   |     |
|----|-----|---|-----|
| Qy | 241 | VHTECCHGDLLECADRADLAKYICENQDSISSKLECCERPLLEKSHCIAEVENDEMPA    | 300 |
| Db | 241 | VHTECCHGDLLECADRADLAKYICENQDSISSKLECCERPLLEKSHCIAEVENDEMPA    | 300 |
| Qy | 301 | DLPSLAADFVESKDCNVAEAKDVLGMLFYEARHPDYSVLLLELAKTYETTTLEK        | 360 |
| Db | 301 | DLPSLAADFVESKDCNVAEAKDVLGMLFYEARHPDYSVLLLELAKTYETTTLEK        | 360 |
| Qy | 361 | CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELPQLEGEYKFNQVALLVRYTKKVPQVST | 420 |
| Db | 361 | CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELPQLEGEYKFNQVALLVRYTKKVPQVST | 420 |
| Qy | 421 | PTLVEVSRNLGKVGSKCKKHPKAKMPCAEEDVLSVVLNQLCVLHEKTPVSDRVTCKCTES  | 480 |
| Db | 421 | PTLVEVSRNLGKVGSKCKKHPKAKMPCAEEDVLSVVLNQLCVLHEKTPVSDRVTCKCTES  | 480 |
| Qy | 481 | LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVKKPKAT  | 540 |
| Db | 481 | LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVKKPKAT  | 540 |
| Qy | 541 | KEQLKAVMDDFAAAFVEKCKKADDDKETCPAECKKLVAASQAALGL                | 585 |
| Db | 541 | KEQLKAVMDDFAAAFVEKCKKADDDKETCPAECKKLVAASQAALGL                | 585 |

Search completed: April 19, 2004, 12:54:56  
Job time : 247.316 secs

Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCYAKVDFEFKPLVBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAAADPHCYAKVDFEFKPLVBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Qy 481 LVNRRPFCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
Db 481 LVNRRPFCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
Qy 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585

## RESULT 14

US-10-319-263-1  
; Sequence 1, Application US/10319263  
; Publication No. US20030180820A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/319,263  
; CURRENT FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-319-263-1

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 4.2e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKRPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKRPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETTLK 360

Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCYAKVDFEFKPLVBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAAADPHCYAKVDFEFKPLVBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Qy 481 LVNRRPFCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
Db 481 LVNRRPFCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
Qy 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585

## RESULT 15

US-10-319-263-2  
; Sequence 2, Application US/10319263  
; Publication No. US20030180820A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/319,263  
; CURRENT FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(585)  
; OTHER INFORMATION: ACETYLYATION  
US-10-319-263-2

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 4.2e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240



```
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Qy 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585

RESULT 12
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 3103; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Qy 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
```

```
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Qy 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585

RESULT 13
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTJK 240
Qy 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
```

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCHIAEVNDEMPA 300  
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCHIAEVNDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
QY 361 CAAADPHECVAKVDFEFKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
DB 361 CAAADPHECVAKVDFEFKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
RESULT 10  
US-10-424-999-11  
; Sequence 11, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Fusion protein human abrogen  
US-10-424-999-11  
Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 4.2e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCHIAEVNDEMPA 300  
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLEKSHCHIAEVNDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
QY 361 CAAADPHECVAKVDFEFKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
DB 361 CAAADPHECVAKVDFEFKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585  
RESULT 11  
US-10-425-000-31  
; Sequence 31, Application US/10425000  
; Publication No. US20040052777A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit  
; TITLE OF INVENTION: Angiogenesis  
; FILE REFERENCE: ST01027-B  
; CURRENT APPLICATION NUMBER: US/10/425,000  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 31  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human derived fusion protein  
US-10-425-000-31  
Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 4.2e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

US-09-832-501-18

ORGANISM: Homo Sapiens

Query Match 100.0%; Score 3103; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 4.2e-269;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGENTFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Db 1 DAHSEVAHRFKDLGENTFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Qy 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAADKAAACLLP 180

Db 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAADKAAACLLP 180

Qy 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240

Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240

Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300

Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300

Qy 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLEKC 360

Db 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLEKC 360

Qy 361 CAADPHCEYKAKVDFEFKPLVEBPQNLIKONCELFEOQGEYKFQNALVRYTKVPOVST 420

Db 361 CAADPHCEYKAKVDFEFKPLVEBPQNLIKONCELFEOQGEYKFQNALVRYTKVPOVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540

Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540

Qy 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

Db 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 9

US-09-833-245-18

Sequence 18, Application US/09833245

Publication No. US20040010134A1

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF546PCT

CURRENT APPLICATION NUMBER: US/09/833,245

PRIOR FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: 60/199,384

PRIOR FILING DATE: 2000-04-25

NUMBER OF SEQ ID NOS: 2267

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 18

TYPE: PRT

LENGTH: 585

ORGANISM: Homo Sapiens

US-09-833-245-18

Query Match 100.0%; Score 3103; DB 11; Length 585;

Best Local Similarity 100.0%; Pred. No. 4.2e-269;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-832-501-18

ORGANISM: Homo Sapiens

Query Match 100.0%; Score 3103; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 4.2e-269;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGENTFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Db 1 DAHSEVAHRFKDLGENTFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Qy 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAADKAAACLLP 180

Db 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAADKAAACLLP 180

Qy 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240

Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240

Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300

Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300

Qy 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLEKC 360

Db 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLEKC 360

Qy 361 CAADPHCEYKAKVDFEFKPLVEBPQNLIKONCELFEOQGEYKFQNALVRYTKVPOVST 420

Db 361 CAADPHCEYKAKVDFEFKPLVEBPQNLIKONCELFEOQGEYKFQNALVRYTKVPOVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540

Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGKPKAT 540

Qy 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

Db 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 8

US-09-833-118-18

Sequence 18, Application US/09833118

Publication No. US20030219875A1

GENERAL INFORMATION:

APPLICANT: Rosen, Craig A.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF544

CURRENT APPLICATION NUMBER: US/09/833,118

PRIOR FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: 60/199,384

PRIOR FILING DATE: 2000-04-25

NUMBER OF SEQ ID NOS: 35

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 18

TYPE: PRT

LENGTH: 585

ORGANISM: Homo Sapiens

US-09-833-118-18

Query Match 100.0%; Score 3103; DB 11; Length 585;

Best Local Similarity 100.0%; Pred. No. 4.2e-269;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18.

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLRLAKTYETTLTK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLRLAKTYETTLTK 360
QY 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540
DB 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 6
US-09-322-322-445
; Sequence 445, Application US/0932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/322,322
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLRLAKTYETTLTK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLRLAKTYETTLTK 360
QY 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540
DB 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US2003019043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS42
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
```

FILING DATE: 19-DEC-1996

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-984-010-26

Query Match

Best Local Similarity 100.0%; Score 3103; DB 10; Length 585;

Mismatches 0; Indels 0; Gaps 0;

Matches 585; Conservative 0;

QY

1

DAKSEVAHRFKDGLGNEFKALVLIAPAFYQLOCPEDHVKLVNEVTEFAKTCVADESAAE

60

DB

1

DAKSEVAHRFKDGLGNEFKALVLIAPAFYQLOCPEDHVKLVNEVTEFAKTCVADESAAE

60

QY

61

NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

DB

61

NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

QY

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

DB

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

QY

181

KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

DB

181

KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

QY

241

VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICAEVENDEMPA

300

DB

241

VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICAEVENDEMPA

300

QY

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

DB

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

QY

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGKYPQNALLVRYTKVPQVST

420

DB

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGKYPQNALLVRYTKVPQVST

420

QY

421

PTLVEVSRLNGKVGSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES

480

DB

421

PTLVEVSRLNGKVGSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES

480

QY

481

LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKKPAT

540

DB

481

LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKKPAT

540

QY

541

KEQLKAVMDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

DB

541

KEQLKAVMDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

RESULT 4

US-09-833-041-18

Sequence 18, Application US/09833041

Publication No. US20030125247A1

GENERAL INFORMATION:

APPLICANT: Haseltine, William A.

APPLICANT: Rosen, Craig A.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF545

CURRENT APPLICATION NUMBER: US/09/833,041

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

FILING DATE: 19-DEC-1996

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-984-010-26

Query Match

Best Local Similarity 100.0%; Score 3103; DB 10; Length 585;

Mismatches 0; Indels 0; Gaps 0;

Matches 585; Conservative 0;

QY

1

DAKSEVAHRFKDGLGNEFKALVLIAPAFYQLOCPEDHVKLVNEVTEFAKTCVADESAAE

60

DB

1

DAKSEVAHRFKDGLGNEFKALVLIAPAFYQLOCPEDHVKLVNEVTEFAKTCVADESAAE

60

QY

61

NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

DB

61

NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

QY

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

DB

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

QY

181

KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

DB

181

KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

QY

241

VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICAEVENDEMPA

300

DB

241

VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICAEVENDEMPA

300

QY

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

DB

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

QY

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGKYPQNALLVRYTKVPQVST

420

DB

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGKYPQNALLVRYTKVPQVST

420

QY

421

PTLVEVSRLNGKVGSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES

480

DB

421

PTLVEVSRLNGKVGSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES

480

QY

481

LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKKPAT

540

DB

481

LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKKPAT

540

QY

541

KEQLKAVMDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

DB

541

KEQLKAVMDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

RESULT 5

US-09-833-117-18

Sequence 18, Application US/09833117

Publication No. US20030171267A1

GENERAL INFORMATION:

APPLICANT: Rosen, Craig A.

APPLICANT: Sadeghi, Homa

APPLICANT: Prior, Christopher P.

APPLICANT: Turner, Andrew J.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF543

CURRENT APPLICATION NUMBER: US/09/833,117

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: 60/199,384

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 3103; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDGLGRENFKALVLIAPAOVLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRRFKDGLGRENFKALVLIAPAOVLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVNCTAPHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVNCTAPHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKHCIAEVNDEM 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKHCIAEVNDEM 300
QY 301 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCNKYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAADPHCEYAKVDFDEPKLVPEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFDEPKLVPEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: April 19, 2004, 12:00:25 ; Search time 246.316 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18  
Perfect score: 3103  
Sequence: 1 DAHSEVAHRFDLGEENFK.....TCFAEGKKLVASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PTC\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PTC\_US\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 3103  | 100.0       | 585    | 9  | US-09-929-552-2   |
| 2          | 3103  | 100.0       | 585    | 10 | US-09-932-613-445 |
| 3          | 3103  | 100.0       | 585    | 10 | US-09-984-010-26  |
| 4          | 3103  | 100.0       | 585    | 10 | US-09-833-041-18  |
| 5          | 3103  | 100.0       | 585    | 10 | US-09-833-117-18  |
| 6          | 3103  | 100.0       | 585    | 10 | US-09-932-322-445 |
| 7          | 3103  | 100.0       | 585    | 10 | US-09-832-501-18  |
| 8          | 3103  | 100.0       | 585    | 11 | US-09-833-118-18  |
| 9          | 3103  | 100.0       | 585    | 11 | US-09-833-245-18  |
| 10         | 3103  | 100.0       | 585    | 12 | US-10-424-999-11  |
| 11         | 3103  | 100.0       | 585    | 12 | US-10-425-000-31  |
| 12         | 3103  | 100.0       | 585    | 13 | US-10-153-064-5   |
| 13         | 3103  | 100.0       | 585    | 14 | US-10-153-604A-5  |
| 14         | 3103  | 100.0       | 585    | 14 | US-10-319-263-1   |
| 15         | 3103  | 100.0       | 585    | 14 | US-10-319-263-2   |

|    |      |       |     |    |                    |                   |
|----|------|-------|-----|----|--------------------|-------------------|
| 16 | 3103 | 100.0 | 585 | 14 | US-10-414-469-1    | Sequence 1, Appli |
| 17 | 3103 | 100.0 | 585 | 14 | US-10-414-469-2    | Sequence 2, Appli |
| 18 | 3103 | 100.0 | 585 | 14 | US-10-413-831-1    | Sequence 1, Appli |
| 19 | 3103 | 100.0 | 585 | 14 | US-10-413-831-2    | Sequence 2, Appli |
| 20 | 3103 | 100.0 | 585 | 15 | US-10-413-832-1    | Sequence 1, Appli |
| 21 | 3103 | 100.0 | 585 | 15 | US-10-413-832-2    | Sequence 2, Appli |
| 22 | 3103 | 100.0 | 585 | 15 | US-10-414-386-1    | Sequence 1, Appli |
| 23 | 3103 | 100.0 | 585 | 15 | US-10-414-386-2    | Sequence 2, Appli |
| 24 | 3103 | 100.0 | 585 | 15 | US-10-233-675A-11  | Sequence 11, Appl |
| 25 | 3103 | 100.0 | 585 | 15 | US-10-462-262-26   | Sequence 26, Appl |
| 26 | 3103 | 100.0 | 604 | 10 | US-09-984-010-7    | Sequence 7, Appli |
| 27 | 3103 | 100.0 | 609 | 10 | US-09-919-039-370  | Sequence 370, App |
| 28 | 3103 | 100.0 | 609 | 13 | US-10-153-064-7    | Sequence 7, Appli |
| 29 | 3103 | 100.0 | 609 | 14 | US-10-153-604A-7   | Sequence 7, Appli |
| 30 | 3103 | 100.0 | 610 | 9  | US-09-984-186-2    | Sequence 23, Appl |
| 31 | 3103 | 100.0 | 610 | 14 | US-10-237-667-2    | Sequence 2, Appli |
| 32 | 3103 | 100.0 | 610 | 14 | US-10-237-708-2    | Sequence 2, Appli |
| 33 | 3103 | 100.0 | 610 | 14 | US-10-237-866-2    | Sequence 2, Appli |
| 34 | 3103 | 100.0 | 610 | 14 | US-10-237-871-2    | Sequence 2, Appli |
| 35 | 3103 | 100.0 | 610 | 14 | US-10-237-871-2    | Sequence 2, Appli |
| 36 | 3103 | 100.0 | 610 | 14 | US-10-237-624-2    | Sequence 13, Appl |
| 37 | 3103 | 100.0 | 616 | 12 | US-10-433-108-13   | Sequence 16, Appl |
| 38 | 3103 | 100.0 | 624 | 12 | US-10-433-108-16   | Sequence 14, Appl |
| 39 | 3103 | 100.0 | 631 | 12 | US-10-433-108-14   | Sequence 15, Appl |
| 40 | 3103 | 100.0 | 640 | 12 | US-10-433-108-15   | Sequence 17, Appl |
| 41 | 3103 | 100.0 | 640 | 12 | US-10-433-108-17   | Sequence 133, App |
| 42 | 3103 | 100.0 | 651 | 13 | US-10-153-064-133  | Sequence 133, App |
| 43 | 3103 | 100.0 | 651 | 14 | US-10-153-604A-133 | Sequence 132, App |
| 44 | 3103 | 100.0 | 652 | 13 | US-10-153-064-132  | Sequence 132, App |
| 45 | 3103 | 100.0 | 652 | 14 | US-10-153-604A-132 | Sequence 132, App |

ALIGNMENTS

RESULT 1  
US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; ; Scto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids

QY 421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEDYLSVTLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEDYLSVTLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRPFCFSALEVDVETVVPKEFNAETFFPHADICTLSEKERQIKQTALVELVVKHKPKAT 540  
DB 481 LVNRPFCFSALEVDVETVVPKEFNAETFFPHADICTLSEKERQIKQTALVELVVKHKPKAT 540  
QY 541 KEQLKAVMDPFAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDPFAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585

RESULT 15  
ABJ00986  
ID ABJ00986 standard; protein; 585 AA.  
XX AC ABJ00986;  
DT 05-SEP-2002 (first entry)  
XX B lymphocyte stimulator protein binding peptide related protein.  
DE B lymphocyte stimulator protein binding protein; BlyS; immune disease;  
KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
KW immunosuppressive; nephrotropic; antithrombotic; antiarthritic;  
KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;  
KW antiasthmatic; antiallergic; thyromimetic; antianaemic; haemostatic;  
KW dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;  
KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
XX Homo sapiens.

XX OS  
XX PN W200216411-A2.  
XX PD 28-FEB-2002.  
XX PF 17-AUG-2001; 2001WO-US025850.  
XX PR 18-AUG-2000; 2000US-0226700P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Beltzer JP, Potter DM, Fleming TL, Rosen CA;  
XX DR WPI; 2002-499775/53.  
XX PT The treatment of various diseases e.g. rheumatoid arthritis, comprises  
PT administering B Lymphocyte stimulator binding polypeptide.  
XX PS Disclosure; Page 379-382; 387pp; English.

XX The present invention relates to the treatment, prevention or  
CC amelioration of a disease or disorder associated with: aberrant B  
CC lymphocyte stimulator (BlyS), BlyS receptor expression or activity; cells  
CC of haematopoietic origin; or proliferative disease; and reducing  
CC inhibiting or stimulating immunoglobulin production; B cell proliferation  
CC and graft rejection involving administration of BlyS binding polypeptide.  
CC The BlyS binding polypeptides are used in the treatment, prevention or  
CC amelioration of diseases such as immune system diseases, proliferative  
CC diseases, diseases of cells of hematopoietic origin, graft rejection,  
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,  
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and  
CC neurodegenerative diseases. The present sequence is a protein described  
CC in the invention

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAB 60  
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFFKAEFVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFFKAEFVSKLVTDLTK 240  
QY 241 VHTCCGHDLLLECADDRADLAKYICENQDISSSKLECCCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLLECADDRADLAKYICENQDISSSKLECCCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKVYAEAKDVFLGMFLYFYARRHPDYSVVLRLRAKTYETTELEK 360  
DB 301 DLPSLAADFVESKDVCKVYAEAKDVFLGMFLYFYARRHPDYSVVLRLRAKTYETTELEK 360  
QY 361 CAAADPHECVAKVPDEFKPLVBPQNLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 420  
DB 361 CAAADPHECVAKVPDEFKPLVBPQNLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEDYLSVTLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEDYLSVTLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRPFCFSALEVDVETVVPKEFNAETFFPHADICTLSEKERQIKQTALVELVVKHKPKAT 540  
DB 481 LVNRPFCFSALEVDVETVVPKEFNAETFFPHADICTLSEKERQIKQTALVELVVKHKPKAT 540  
QY 541 KEQLKAVMDPFAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDPFAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585

Search completed: April 19, 2004, 11:51:14  
Job time : 348.086 secs



Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGBENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHRFKDGBENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120

QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 DB 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240  
 DB 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
 DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360  
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
 DB 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCCHPKAKMPCAEDYLSVNLQCVLHEKTPVSDRVTCCCTES 480  
 DB 421 PTLVEVSRNLGKVGSKCCHPKAKMPCAEDYLSVNLQCVLHEKTPVSDRVTCCCTES 480

QY 481 LVNRRFCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPAT 540  
 DB 481 LVNRRFCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPAT 540

QY 541 KEQLKAVMDDFAAFVKECKCADDKETCFABEGKGLVAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFVKECKCADDKETCFABEGKGLVAASQAALGL 585

## RESULT 14

AAU75220  
 ID AAU75220 standard; protein; 585 AA.

XX AAU75220;

XX 21-MAY-2002 (first entry)

XX Mature form of human serum albumin (HSA or HA).

XX Albumin fusion protein; therapeutic protein; immune disorder;  
 KW autoimmune disorder; blood-related disorder; hyperproliferative disorder;  
 KW renal disorder; cardiovascular disorder; respiratory disorder;  
 KW neurological disorder; endocrine disorder; reproductive system disorder;  
 KW gastrointestinal disorder; infectious disease; wound healing;  
 KW human serum albumin; HSA; HA.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Protein 1..585  
 FT /label= Mature\_HSA

XX WO200179271-A1.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US012009.

XX 12-APR-2000; 2000US-0229358P.

PR 25-APR-2000; 2000US-0199384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 PA (PRIN-) PRINCIPIA PHARM CORP.  
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
 XX Ballance DJ, Sleep D, Turner AJ, Sadeghi H, Prior CP;  
 XX WPI; 2002-179329/23.  
 DR N-PSDB; ABK13862.  
 XX New albumin fusion proteins with extended shelf life, useful for treating  
 XX leukemia, warts, hepatitis, multiple sclerosis and AIDS, comprises  
 XX therapeutic protein fused to albumin.  
 XX Claim 1; Fig 15; 338pp; English.  
 XX The present invention relates to albumin fusion proteins comprising a  
 XX therapeutic protein and albumin (e.g. human serum albumin (HSA or HA)).  
 XX The albumin fusion proteins are useful for treating, preventing, or  
 XX ameliorating various disorders. Such disorders include immune disorders,  
 XX autoimmune disorders, blood-related disorders, hyperproliferative  
 XX disorders, renal disorders, cardiovascular disorders, respiratory  
 XX disorders, neurological disorders, endocrine disorders, reproductive  
 XX system disorders, gastrointestinal disorders, infectious disease, and  
 XX wound healing. Therapeutic proteins can be stabilised to extend shelf  
 XX life and/or retain the protein's activity for extended periods of time in  
 XX solution, in vivo or in vitro by genetically or chemically fusing the  
 XX protein to albumin or its fragment or variant. In addition the use of  
 XX albumin fusion proteins reduces the need to formulate protein solutions  
 XX with large excesses of carrier proteins to prevent loss of therapeutic  
 XX protein due to factors such as binding to the container. The extension of  
 XX shelf life was tested by measuring biological activity (Nb2 cell  
 XX proliferation) of human albumin-human growth hormone (HA-hGH) fusion  
 XX protein remaining after incubation in cell culture media for up to 3  
 XX weeks. At week 3 there was still approximately 95% cell proliferation  
 XX compared to no activity of unfused hGH. The present sequence represents  
 XX the mature form of HSA which can be used to produce the albumin fusion  
 XX proteins of the invention

Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKOLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60

DB 1 DAHSEVAHRFKOLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120

QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

DB 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

DB 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360

DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420

DB 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420

|           |  |     |
|-----------|--|-----|
| 191       | KLDELURDEGKASSAKQRLUKCASLQKFGERAFAKANAVASLSORFFKAEPFAEVSGLVTDLT  | 360 |
| 241       | VHTECHGDLLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA   | 300 |
| 241       | VHTECHGDLLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA   | 300 |
| 301       | DLPSLAADFVSKQVCKYAAKQVFLGMFLYAYARRHPDYSVLLLLRLAKTYETTLK  | 360 |
| 301       | DLPSLAADFVSKQVCKYAAKQVFLGMFLYAYARRHPDYSVLLLLRLAKTYETTLK  | 360 |
| 361       | CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST  | 420 |
| 361       | CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST  | 420 |
| 421       | PTLVEVSRLGKVGSKCKHPKAMPKQVCAEDYLSVTLNQLCVLHEKTPVSDRVTCKCTES  | 480 |
| 421       | PTLVEVSRLGKVGSKCKHPKAMPKQVCAEDYLSVTLNQLCVLHEKTPVSDRVTCKCTES  | 480 |
| 481       | LVNRRPCFSALEVDITYYPKFNATFTFHADICTLSEKEROIKQTALVELVGHKPKAT  | 540 |
| 481       | LVNRRPCFSALEVDITYYPKFNATFTFHADICTLSEKEROIKQTALVELVGHKPKAT  | 540 |
| 541       | KEQLKAVMDDFAAFVEKCKCKADDKTCFAEKGKKLVAASQAALGL  | 585 |
| 541       | KEQLKAVMDDFAAFVEKCKCKADDKTCFAEKGKKLVAASQAALGL  | 585 |
| RESULT 13 |  |     |
| AAE08578  |  |     |
| ID        | AAE08578 standard; protein; 585 AA.  |     |
| XX        | AAE08578;  |     |
| XX        | 19-NOV-2001 (first entry)  |     |
| XX        | Human serum albumin (HSA).   |     |
| XX        | Human; albumin; cancer; cell proliferation; drug screening; biopsy.  |     |
| XX        | Homo sapiens.  |     |
| XX        | US6274305-B1.  |     |
| XX        | 14-AUG-2001.   |     |
| XX        | 19-DEC-1996; 96US-00769746.  |     |
| XX        | 19-DEC-1996; 96US-00769746.  |     |
| XX        | (TUFT ) UNIV TUFTS.  |     |
| XX        | Sonnenschein C, Soto AM;   |     |
| XX        | WPI: 2001-540371/60.   |     |
| XX        | N-PSDB; AAD11488.  |     |
| XX        | Measuring human cell proliferation, useful in drug screening to determine the potential for inhibiting cancer cell proliferation and for evaluating biopsied tumors, comprises employing albumin-derived peptide.  |     |
| XX        | Claim 1; Fig 1; 20pp; English.   |     |
| XX        | The invention related to a method for testing cancer cells. The method is useful for measuring human cancer cell proliferation, particularly for determining the potential for inhibiting cancer cells proliferation using albumin-derived peptides. The invention is also useful for drug screening assays, as well as for evaluating biopsied tumors. The present sequence is human serum albumin (HSA) related to the invention |     |
| XX        | Sequence 585 AA;   |     |
| XX        | Query Match 100.0%; Score 3103; DB 4; Length 585;  |     |
| XX        | Best Local Similarity 100.0%; Pred. NO. 1.5e-254;  |     |

FT Domain /label= Loop\_IX  
 FT 461..475  
 FT /label= Loop\_X  
 FT 478..486  
 FT /label= Loop\_XI  
 FT 560..566  
 FT /label= Loop\_XII  
 PN WO200179258-A1.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US012008.  
 XX  
 PR 12-APR-2000; 2000US-0229359P.  
 PR 25-APR-2000; 2000US-019384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (PRIN-) PRINCIPIA PHARM CORP.  
 XX  
 PI Rosen CA, Sadeghi H, Prior CP, Turner AJ;  
 XX  
 XX WPI; 2001-602931/68.  
 DR N-PSDB; AAD22287.  
 DR  
 XX Albumin fusion proteins comprising a therapeutic protein and albumin,  
 PT useful in the treating metastatic renal cell carcinoma, metastatic  
 PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human  
 PT immunodeficiency virus) or infection.  
 XX  
 XX Claim 1; Fig 9; 325pp; English.  
 XX  
 XX The invention relates to albumin fusion proteins comprising therapeutic  
 CC protein and human albumin (HA). The albumin fusion proteins are useful in  
 CC the treatment, prevention, diagnosis, and/or detection of diseases,  
 CC disorders such as immune system disorders (transplant rejection); blood  
 CC related disorders (myocardial infarction); hyperproliferative disorders  
 CC (childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);  
 CC cardiovascular disorders (arrhythmias); respiratory disorders (non-  
 CC allergic rhinitis); neurological diseases (Alzheimer's disease);  
 CC endocrine disorders (pheochromocytoma); reproductive system disorders  
 CC (syphilis); infectious diseases (measles); gastrointestinal disorders  
 CC (irritable bowel syndrome) and wound healing. The albumin fusion proteins  
 CC are also used in the treatment of metastatic renal cell carcinoma,  
 CC metastatic melanoma, malignant melanoma and HIV (human immunodeficiency  
 CC virus) infection. Nucleic acid encoding albumin fusion protein is useful  
 CC in gene therapy. The present sequence is human albumin (HA) protein  
 XX  
 XX Sequence 585 AA;  
 SQ  
 Query Match 100.0%; Score 3103; DB 4; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHKSVAHRFKDLGEENFKALVIAFAQVLOCPEDHVKLVNEVTEPAKTCVADESAG 60  
 DB 1 DAHKSVAHRFKDLGEENFKALVIAFAQVLOCPEDHVKLVNEVTEPAKTCVADESAG 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120  
 QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCCAADKACLLP 180  
 DB 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCCAADKACLLP 180  
 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPPKAEFAVSKLVTDLTK 240  
 DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPPKAEFAVSKLVTDLTK 240  
 QY 241 VHTCCGHDLLFCADDEADLAKYICENQDSISSKLKCECKEFLLEKSHCIAEVENDEMPA 300

Db 241 VHTCCGHDLLFCADDEADLAKYICENQDSISSKLKCECKEFLLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVLLLRLLAKTYETLEK 360  
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVLLLRLLAKTYETLEK 360  
 QY 361 CAAADPHECVAKVDFDFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKVPQVST 420  
 Db 361 CAAADPHECVAKVDFDFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKVPQVST 420  
 QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480  
 Db 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
 Db 481 LVNRRPCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
 QY 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFPAEEGKKLVAAASQAALGL 585  
 Db 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFPAEEGKKLVAAASQAALGL 585  
 RESULT 12  
 ABB79006  
 ID ABB79006 standard; protein; 585 AA.  
 XX  
 AC ABB79006;  
 XX  
 DT 01-AUG-2002 (first entry)  
 XX  
 DE Human mature albumin protein SEQ ID NO:18.  
 XX  
 KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;  
 KW albumin fusion protein; cytostatic; anorectic; immunosuppressive;  
 KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;  
 KW non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;  
 KW type I diabetes mellitus; rheumatoid arthritis.  
 XX  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 1..194  
 FT /label= 1  
 FT Domain 1..105  
 FT /label= subdomain  
 FT Disulfide-bond 53..62  
 FT Disulfide-bond 75..91  
 FT Disulfide-bond 90..101  
 FT Region 106..119  
 FT /note= "flexible inter-subdomain linker region"  
 FT Domain 120..194  
 FT /label= subdomain  
 FT Disulfide-bond 124..169  
 FT Disulfide-bond 168..177  
 FT Domain 195..387  
 FT /label= 2  
 FT Domain 195..291  
 FT /label= subdomain  
 FT Disulfide-bond 245..253  
 FT Disulfide-bond 265..279  
 FT Disulfide-bond 278..289  
 FT Region 292..315  
 FT /note= "flexible inter-subdomain linker region"  
 FT Domain 316..387  
 FT /label= subdomain  
 FT Disulfide-bond 316..361  
 FT Disulfide-bond 360..369  
 FT Domain 388..585  
 FT /label= 3  
 FT Domain 388..491  
 FT /label= subdomain  
 FT Disulfide-bond 392..438

FT Domain /label= Loop\_III 170..176  
FT /label= Loop\_IV 247..252  
FT Domain /label= Loop\_V 266..277  
FT /label= Loop\_VI 280..288  
FT Domain /label= Loop\_VII 362..368  
FT /label= Loop\_VIII 439..447  
FT Domain /label= Loop\_IX 461..475  
FT /label= Loop\_X 478..485  
FT Domain /label= Loop\_XI 560..566  
FT /label= Loop\_XII  
WC200179443-A2.  
25-OCT-2001.  
12-APR-2001; 2001WO-US011924.  
12-APR-2000; 2000US-0229358P.  
25-APR-2000; 2000US-0199384P.  
21-DEC-2000; 2000US-0256931P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Haseltine WA;  
XX WPI; 2001-616754/71.  
DR N-PSDB; AAD21639.  
XX  
XX Albumin fusion proteins comprising a therapeutic protein and albumin,  
PT useful in the treating immune system disorders (e.g. transplant  
PT rejection), blood related disorders (e.g. myocardial infarction) and  
PT hyperproliferative disorders.  
XX  
XX Claim 1; Fig 9; 380pp; English.  
PS  
XX The invention relates to albumin fusion proteins comprising therapeutic  
CC protein and human albumin (HA). Therapeutic protein fused to albumin have  
CC an extended shelf-life. The albumin fusion proteins are useful in the  
CC treatment, prevention, diagnosis and/or detection of diseases, disorders  
CC such as immune system disorders (e.g. transplant rejection), blood  
CC related disorders (e.g. myocardial infarction), hyperproliferative  
CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders (e.g.  
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),  
CC respiratory disorders (e.g. non-allergic rhinitis), neurological diseases  
CC (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),  
CC reproductive system disorders (e.g. syphilis), infectious diseases (e.g.  
CC measles), gastrointestinal disorders (e.g. irritable bowel syndrome) and  
CC wound healing. Nucleic acids encoding albumin fusion protein is used in  
CC gene therapy. The present sequence is human albumin (HA) protein  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRPKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHSEVAHRPKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Qy 121 DVNCTAFHDNEETFLKKLYLIEIARRHPYFYAPPELLFFFAKRYKAAFTCCQAAADKACLLP 180  
Db 121 DVNCTAFHDNEETFLKKLYLIEIARRHPYFYAPPELLFFFAKRYKAAFTCCQAAADKACLLP 180  
Qy 181 KLDLDEGKASAKORLKCASLQKFGERAFAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDLDEGKASAKORLKCASLQKFGERAFAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Qy 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDEMPA 300  
Db 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLEKC 360  
Qy 361 CAAADPHECYAKVDFEKFPLVEBPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420  
Db 361 CAAADPHECYAKVDFEKFPLVEBPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480  
Qy 481 LVNERPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540  
Db 481 LVNERPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCCCKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCCCKADDKETCFABEGKKLVAASQAALGL 585  
RESULT 11  
AAE13399  
ID AAE13399 standard; protein; 585 AA.  
XX  
AC AAE13399;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE Human albumin (HA) protein.  
XX  
KW Human; albumin; HA; fusion protein; immune system disorder; syphilis;  
KW transplant rejection; blood related disorder; myocardial infarction;  
KW hyperproliferative disorder; acute myeloid leukaemia; renal disorder;  
KW glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;  
KW respiratory disorder; neurological disease; Alzheimer's disease;  
KW endocrine disorder; pheochromocytoma; reproductive system disorder;  
KW measles; gastrointestinal disorder; irritable bowel syndrome; HIV;  
KW human immunodeficiency virus; wound healing; renal cell carcinoma;  
KW melanoma; gene therapy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Domain /label= Loop\_I 54..61  
FT Domain /label= Loop\_II 76..89  
FT Domain /label= Loop\_III 92..100  
FT Domain /label= Loop\_IV 170..176  
FT Domain /label= Loop\_V 247..252  
FT Domain /label= Loop\_VI 266..277  
FT Domain /label= Loop\_VII 280..288  
FT Domain /label= Loop\_VIII 362..368  
FT Domain /label= Loop\_VIII 439..447



QY 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRPCFSALEVDVETVPKFNABETTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
DB 481 LVNRPCFSALEVDVETVPKFNABETTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585

## RESULT 8

AA52567  
ID AAM52567 standard; protein; 585 AA.

AC AAM52567;

DT 05-FEB-2002 (first entry)

XX Mature human serum albumin.

XX Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;

XX nontropic; neuroprotective; gene therapy; immune disorder; wound healing;

XX hyperproliferative disorder; renal disorder; cardiovascular disorder;

XX respiratory disorder; neurological disease; endocrine disorder;

XX reproductive system disorder; infectious disease;

XX gastrointestinal disorder.

XX Homo sapiens.

OS WO200179444-A2.

FN 25-OCT-2001.

PD 12-APR-2001; 2001WO-US012013.

PP 12-APR-2000; 2000US-0229358P.

PR 21-DEC-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Haseltine WA;

XX WPI; 2001-61675/71.

DR N-PSDB; ABA03057.

XX Albumin fusion proteins comprising a therapeutic protein and albumin,

XX useful in the treating immune system disorders (e.g. transplant

XX rejection), blood related disorders (e.g. myocardial infarction) and

XX hyperproliferative disorders.

XX Claim 1; Fig 15; 606pp; English.

PS The present invention relates to albumin fusion proteins, which comprise

XX a therapeutic protein and albumin. The present sequence is the protein

XX sequence for mature human serum albumin (HA), which was used to generate

XX the fusion proteins of the present invention. The albumin fusion proteins

XX are useful in the treatment, prevention, diagnosis, and/or detection of

XX diseases/disorders such as immune system disorders (e.g. transplant

XX rejection), blood related disorders (e.g. myocardial infarction),

XX hyperproliferative disorders (e.g. childhood acute myeloid leukemia),

XX renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.

XX arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),

XX neurological diseases (e.g. Alzheimer's disease), endocrine disorders

XX (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),

CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.  
CC irritable bowel syndrome) and wound healing  
XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVHRFVDCLEENFKALVLAFAQYLQCCPFEDHVKLVNVEVTEPAKTCVADSSAE 60  
DB 1 DAHSEVHRFVDCLEENFKALVLAFAQYLQCCPFEDHVKLVNVEVTEPAKTCVADSSAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGPERNECFLOKDDNPMLPLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGPERNECFLOKDDNPMLPLVRPEV 120  
QY 121 DVNCTAFHNEETFLKKYLYEYARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVNCTAFHNEETFLKKYLYEYARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAMAVARLSORFPKAEPAFVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLOKGERAFKAMAVARLSORFPKAEPAFVSKLVTDLTK 240  
QY 241 VHTTECHGDLLECCADDDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTTECHGDLLECCADDDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKVCKNVAEAKDVFLGMFLYEYARRHPDYVSVLLLLAKTYETTLK 360  
DB 301 DLPSLAADFVSKVCKNVAEAKDVFLGMFLYEYARRHPDYVSVLLLLAKTYETTLK 360  
QY 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRPCFSALEVDVETVPKFNABETTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
DB 481 LVNRPCFSALEVDVETVPKFNABETTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585

## RESULT 9

AAE12403  
ID AAE12403 standard; protein; 585 AA.

AC AAE12403;

XX 18-DEC-2001 (first entry)

DT Human albumin (HA).

XX Human; albumin; HA; immune system disorder; transplant rejection;

XX blood related disorder; myocardial infarction; glomerulonephritis;

XX hyperproliferative disorder; childhood acute myeloid leukaemia;

XX renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;

XX arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;

XX neurological disease; Alzheimer's disease; endocrine disorder; syphilis;

XX pheochromocytoma; reproductive system disorder; neuroprotective; nontropic;

XX infectious disease; gastrointestinal disorder; wound healing; infection;

XX irritable bowel syndrome; HIV; human immunodeficiency virus infection;

XX cytotoxic; antinflammatory; gene therapy; immunosuppressive; cardiant;

XX antiarthritic; antirheumatic; renal disorder; antimicrobial.

XX Homo sapiens.

XX New method for the continuous detection of ischemic states comprises  
 PT detecting and quantifying the existence of an alteration of the serum  
 PT protein albumin.

XX Disclosure; Page 97-100; 105pp; English.

XX The present sequence represents human albumin protein. The specification  
 CC describes a method for the continuous detection of ischemic states. The  
 CC method comprises detecting and quantifying the existence of an alteration  
 CC of the serum protein albumin. The method comprises contacting a  
 CC biological sample containing albumin from the patient with an excess  
 CC quantity of a metal ion salt, where the metal ion binds to the N-terminus  
 CC of naturally occurring human albumin, to form a mixture containing bound  
 CC metal ions and unbound metal ions, and then determining the amount of  
 CC metal ions bound to the albumin N-terminus. The amount of bound metal  
 CC ions is correlated to a known value to determine the occurrence or non-  
 CC occurrence of an ischemic event. The methods are useful for detection of  
 CC ischemic states. The methods are also useful for distinguishing  
 CC peroperative ischemia from ischemia caused by, amongst other things,  
 CC myocardial infarctions and progressive coronary artery disease

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 3; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60

61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
 DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
 DB 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

241 VHTCCGDLLECADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
 DB 241 VHTCCGDLLECADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

301 DLPSLAADFVESKDVCKQYAEAKDVLGMFLFYARRHPDYSVLLLRKAKTYETTTLEKC 360  
 DB 301 DLPSLAADFVESKDVCKQYAEAKDVLGMFLFYARRHPDYSVLLLRKAKTYETTTLEKC 360

361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420  
 DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420

421 PTLVEVSRNLGVGSKCCCHPKAKMPCAEDYLSVLNQLCVLHKETPSDRVTKCCTES 480  
 DB 421 PTLVEVSRNLGVGSKCCCHPKAKMPCAEDYLSVLNQLCVLHKETPSDRVTKCCTES 480

481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVXHKPKAT 540  
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVXHKPKAT 540

541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 7

AY83946  
 ID AY83946 standard; protein; 585 AA.

XX

AA83946;

28-JUL-2000 (first entry)

Yeast codon-biased recombinant human serum albumin protein.

Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
 overlapping oligonucleotide; expression vector.

Homo sapiens.

Synthetic.

CN1239103-A.

22-DEC-1999.

17-JUN-1998; 98CN-00102506.

17-JUN-1998; 98CN-00102506.

(HAJJ-) HALJI BIOENGINEERING CO LTD.

Li S, Lu D;

WPI; 2000-351198/31.

N-PSDB; AAA10091.

Process for preparing recombinant human serum albumin comprising yeast  
 biased sex codons - uses a recombinant DNA technique.

Disclosure; Fig 1; 44pp; Chinese.

The method relates to a method of recombinantly producing human serum  
 albumin (HSA) in yeast by altering the coding sequence of HSA to comprise  
 a yeast codon bias. The complete HSA gene (AAA10091) was generated as  
 three synthetic fragments (AAA10092-Alc094) joined by recombinant DNA  
 technology. Each HSA fragment was synthesised from overlapping  
 oligonucleotide fragments that were extended. This sequence represents  
 the complete sequence of the HSA encoded by the human gene with a yeast  
 codon bias. The invention also covers a recombinant expression vector,  
 yeast host cells carrying the recombinant expression vector and the  
 process for producing human serum albumin in the yeast host cell,  
 especially in secretory mode

Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 3; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60

DB 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60

61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

DB 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

241 VHTCCGDLLECADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

DB 241 VHTCCGDLLECADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

301 DLPSLAADFVESKDVCKQYAEAKDVLGMFLFYARRHPDYSVLLLRKAKTYETTTLEKC 360

DB 301 DLPSLAADFVESKDVCKQYAEAKDVLGMFLFYARRHPDYSVLLLRKAKTYETTTLEKC 360

QY 301 DLPSSAADFVESKDVCKYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSSAADFVESKDVCKYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALIVRTKKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALIVRTKKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHKTPTVSDRVTKCCTES 480  
DB 421 PTLVEVSRLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHKTPTVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDYETVVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 5  
AAO20111  
ID AAO20111 standard; protein; 585 AA.  
AC AAO20111;  
XX  
XX  
DT 06-AUG-2002 (first entry)  
DE HSA protein sequence related to the growth hormone protein.  
XX  
XX  
KW Serum albumin-growth hormone fusion protein; growth hormone;  
KW Down's syndrome.  
XX  
XX  
OS Unidentified.  
XX  
XX  
PN KR99076789-A.  
XX  
XX  
PD 15-OCT-1999.  
XX  
XX  
PF 25-JUN-1998; 98KR-00704914.  
XX  
XX  
PR 30-DEC-1995; 95GB-00026733.  
PR 19-DEC-1996; 96WO-GB003164.  
XX  
XX  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
XX  
PI Ballance DJ;  
DR WPI; 1997-363680/33.  
DR N-PSDB; AAK99568.  
XX  
XX  
PT Serum albumin-growth hormone fusion protein - useful to treat growth  
PT hormone related diseases, e.g. Down's syndrome.  
XX  
XX  
PS Disclosure; Fig 6; 2lpp; Korean.  
XX  
XX  
CC The invention relates to a serum albumin-growth hormone fusion protein -  
CC useful to treat growth hormone related diseases such as Down's syndrome.  
CC This sequence represents a HSA protein related to the serum albumin-  
CC growth hormone protein of the invention  
XX  
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDIDGENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHFKDIDGENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVNCTAFHNEETFLKXLYEYARRHPYFVAPPELLFAFKRYKAAAFTECCQADKAACLLP 180  
DB 121 DVNCTAFHNEETFLKXLYEYARRHPYFVAPPELLFAFKRYKAAAFTECCQADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFPKAEFAVSKLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFPKAEFAVSKLVTDLT 240  
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
DB 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSSAADFVESKDVCKYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSSAADFVESKDVCKYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALIVRTKKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALIVRTKKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHKTPTVSDRVTKCCTES 480  
DB 421 PTLVEVSRLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHKTPTVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDYETVVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 6  
AA84873  
ID AA84873 standard; protein; 585 AA.  
AC AA84873;  
XX  
XX  
DT 08-AUG-2000 (first entry)  
DE Amino acid sequence of a human albumin protein.  
XX  
XX  
KW Human; albumin; ischemic state; serum protein; metal ion salt;  
KW perioperative ischemia; ischemia; myocardial infarction;  
KW progressive coronary artery disease.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "optionally acetylated, and claimed under claim  
FT 56"  
XX  
XX  
PN WO200020840-A1.  
XX  
XX  
PD 13-APR-2000.  
XX  
XX  
PF 01-OCT-1999; 95WO-US022905.  
XX  
XX  
PR 02-OCT-1998; 98US-00165581.  
PR 02-OCT-1998; 98US-00165526.  
PR 02-OCT-1998; 98US-0102738P.  
PR 11-JAN-1999; 99US-0115392P.  
XX  
XX  
PA (ISCH-) ISCHEMIA TECHNOLOGIES INC.  
XX  
XX  
PI Bar-Or D, Lau E, Winkler JV;  
XX  
XX  
DR WPI; 2000-303843/26.



XX JP02227079-A.  
 XX PN  
 XX PD  
 XX PF  
 XX PF 25-AUG-1989; 89JP-00217540.  
 XX PR  
 XX PR 06-OCT-1988; 88JP-00250926.  
 XX (TOFU) TONEN CORP.  
 XX PA  
 XX XX  
 XX WPI; 1990-317325/42.  
 XX DR N-PSDB; AAQ06099.  
 XX PT  
 XX PT New human serum albumin fragments - used to bond to medicines and for stable folding of protein(s).  
 XX XX  
 XX PS Claim 1; Fig 8; 24pp; Japanese.  
 XX XX  
 CC Fragments A-C of HSA are expressed as fusion proteins with the signal peptide of E. coli alkaline phosphatase. The fragments are selected for their specific properties. The C-terminal truncated fragment, B, does not bind long-chain fatty acids but does bind to various medicines at the central region. The N-terminal truncated fragment, C, has good stability in protein folding. The central segment, A, has characteristics of both B and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)  
 XX XX  
 XX SQ Sequence 585 AA;  
 Query Match 100.0%; Score 3103; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
 QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 QY 181 KLDELADDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
 Db 181 KLDELADDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
 QY 241 VHTCCGDLLECADDDRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
 Db 241 VHTCCGDLLECADDDRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADPFVSKVCKVNAEAKOVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 360  
 Db 301 DLPSLAADPFVSKVCKVNAEAKOVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 360  
 QY 361 CAAADPHECVAKVDFEFKPLVEEPONLIKONCELFEOQLGEYKFONALLVRVTKVPOYST 420  
 Db 361 CAAADPHECVAKVDFEFKPLVEEPONLIKONCELFEOQLGEYKFONALLVRVTKVPOYST 420  
 QY 421 PTLVEVSRNLGVKSGCKKHPRKMPCAEDYLSVVLNQLCVLHKETPVSORVTKCTTES 480  
 Db 421 PTLVEVSRNLGVKSGCKKHPRKMPCAEDYLSVVLNQLCVLHKETPVSORVTKCTTES 480  
 QY 481 LVNRRFCFSALEVDETVYVPEKNAETFTPHADICITLSEKEROIKKOTALVELVHKPKAT 540  
 Db 481 LVNRRFCFSALEVDETVYVPEKNAETFTPHADICITLSEKEROIKKOTALVELVHKPKAT 540  
 QY 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFAEKGKLVAAASQAALGL 585

Db 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFAEKGKLVAAASQAALGL 585  
 RESULT 4  
 AAR80301  
 XX AAR80301 standard; protein; 585 AA.  
 XX AC AAR80301;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 17-JAN-1996 (first entry)  
 XX XX  
 XX DE Human serum albumin.  
 XX KW Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.  
 XX OS Homo sapiens.  
 XX PN WO9523857-A1.  
 XX PD 08-SEP-1995.  
 XX PF 01-MAR-1995; 95WO-GB0000434.  
 XX PR 05-MAR-1994; 94GB-00004270.  
 XX XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
 XX PI KerryWilliams SM, Gilbert SC;  
 XX WPI; 1995-320572/41.  
 XX DR N-PSDB; AAQ98695.  
 XX PT Yeast with reduced levels of aspartyl protease 3 proteolytic activity - used to secrete human albumin without prodn. of the 45 kD fragment.  
 XX PS Example 1; Page 26-28; 50pp; English.  
 CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected to site-directed mutagenesis to investigate the role of endoproteases in the generation of a 45 kDa albumin fragment obtd. when the cDNA is expressed in S. cerevisiae. Mutations were: R410A; L407A, L408V, V409A; and R410A, K413Q, K414Q. The latter set of mutations, especially, CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct CC PI field.)  
 XX SQ Sequence 585 AA;  
 Query Match 100.0%; Score 3103; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
 QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 QY 181 KLDELADDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
 Db 181 KLDELADDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
 QY 241 VHTCCGDLLECADDDRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
 Db 241 VHTCCGDLLECADDDRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPA 300

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 1 DAHKEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 QY 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
 DB 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
 QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
 DB 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
 QY 241 VHTCCHGDLLECCADRADLAKYI CENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
 DB 241 VHTCCHGDLLECCADRADLAKYI CENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEKC 360  
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEKC 360  
 QY 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGEYKFONALLVRYTKVPOVST 420  
 DB 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGEYKFONALLVRYTKVPOVST 420  
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540  
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEEGKKLVAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEEGKKLVAASQAALGL 585

RESULT 2

AA05318  
 ID AAR05318 standard; protein; 585 AA.

AC AAR05318;  
 XX 08-OCT-1990 (first entry)  
 DT Human serum albumin gene product.  
 DE Human serum albumin; HSA-A; yeast; ds.

XX Homo sapiens.  
 XX JP02117384-A.  
 XX 01-MAY-1990.  
 XX 26-OCT-1988; 88JJP-00268302.  
 XX 26-OCT-1988; 88JJP-00268302.  
 XX (TOFU) TOA NENRYO KOGYO KK.  
 XX WPI; 1990-176228/23.  
 DR N-PSDB; AAQ04719.

PT Human serum albumin prepn. by yeast host - by culturing transformed  
 PT plasmid yeast to produce serum, and removing it.  
 PS Disclosure; Page ?; -pp; Japanese.

XX

CC Mature HSA-A may be produced using the sequence incorporated into a  
 CC plasmid vector with suitable controllers, and transferred to a yeast  
 CC expression system  
 XX  
 SQ Sequence 585 AA;  
 Query Match 100.0%; Score 3103; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHKEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 1 DAHKEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 QY 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
 DB 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
 QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
 DB 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
 QY 241 VHTCCHGDLLECCADRADLAKYI CENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
 DB 241 VHTCCHGDLLECCADRADLAKYI CENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEKC 360  
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEKC 360  
 QY 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGEYKFONALLVRYTKVPOVST 420  
 DB 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGEYKFONALLVRYTKVPOVST 420  
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540  
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEEGKKLVAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEEGKKLVAASQAALGL 585

RESULT 3

AA08457  
 ID AAR08457 standard; protein; 585 AA.

AC AAR08457;  
 XX 25-MAR-2003 (revised)  
 DT 16-APR-1991 (first entry)  
 DE Human serum albumin.

XX HSA; folding; ss.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 XX 1. .303  
 FT Region /label= B  
 FT 123. .585  
 FT /label= C  
 FT 123. .303  
 FT /label= A

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 344.086 Seconds  
(without alignments)  
480.375 Million cell updates/sec

Title: US-09-832-929-18  
Perfect score: 3103  
Sequence: 1 DAHSEVAHFKVLGHEHFK.....TCFAEBSKKLVAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_29Jan04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1          | 3103  | 100.0       | 585    | 1 AAP90388 | Mature hu   |
| 2          | 3103  | 100.0       | 585    | 2 AAR05318 | Human ser   |
| 3          | 3103  | 100.0       | 585    | 2 AAR08457 | Human ser   |
| 4          | 3103  | 100.0       | 585    | 2 AAR80301 | Human ser   |
| 5          | 3103  | 100.0       | 585    | 2 AAO20111 | HSA prote   |
| 6          | 3103  | 100.0       | 585    | 3 AAY84873 | Amino aci   |
| 7          | 3103  | 100.0       | 585    | 3 AAY83946 | Yeast cod   |
| 8          | 3103  | 100.0       | 585    | 4 AAM52567 | Mature hu   |
| 9          | 3103  | 100.0       | 585    | 4 AAE12403 | Human alb   |
| 10         | 3103  | 100.0       | 585    | 4 AAE13129 | Human alb   |
| 11         | 3103  | 100.0       | 585    | 4 AAE13399 | Human alb   |
| 12         | 3103  | 100.0       | 585    | 4 ABB79006 | Human mat   |
| 13         | 3103  | 100.0       | 585    | 4 AAE08578 | Human ser   |
| 14         | 3103  | 100.0       | 585    | 5 AAU75220 | Mature fo   |
| 15         | 3103  | 100.0       | 585    | 5 ABJ00986 | B lymphoc   |
| 16         | 3103  | 100.0       | 585    | 5 ABG33321 | Human ser   |
| 17         | 3103  | 100.0       | 585    | 5 ABG33847 | Human B L   |
| 18         | 3103  | 100.0       | 585    | 5 ABG71291 | Glycosyla   |
| 19         | 3103  | 100.0       | 585    | 6 ABR55695 | Human alb   |
| 20         | 3103  | 100.0       | 585    | 7 ABR42606 | Human ser   |
| 21         | 3103  | 100.0       | 585    | 7 ADC16767 | Human ser   |
| 22         | 3103  | 100.0       | 585    | 7 ADD06469 | Human ser   |
| 23         | 3103  | 100.0       | 585    | 7 ADD68016 | Mature fo   |
| 24         | 3103  | 100.0       | 609    | 3 AAB36542 | Recombina   |
| 25         | 3103  | 100.0       | 609    | 3 AAY78147 | Pre human   |

|    |      |       |     |            |           |
|----|------|-------|-----|------------|-----------|
| 26 | 3103 | 100.0 | 609 | 3 AAB36549 | Recombina |
| 27 | 3103 | 100.0 | 609 | 6 ABU57252 | Human ser |
| 28 | 3103 | 100.0 | 609 | 6 ABU57253 | Human ser |
| 29 | 3103 | 100.0 | 609 | 7 ADA24217 | Human ser |
| 30 | 3103 | 100.0 | 609 | 7 ADD06471 | Human ser |
| 31 | 3103 | 100.0 | 609 | 8 ADE77205 | Human pro |
| 32 | 3103 | 100.0 | 610 | 2 AAR39510 | Chimeric  |
| 33 | 3103 | 100.0 | 616 | 6 AAE30916 | Val8-GLP- |
| 34 | 3103 | 100.0 | 624 | 6 AAE30919 | Human ser |
| 35 | 3103 | 100.0 | 631 | 6 AAE30917 | Val8-GLP- |
| 36 | 3103 | 100.0 | 640 | 6 AAE30920 | Exendin-4 |
| 37 | 3103 | 100.0 | 640 | 6 AAE30918 | Gly8-Glu2 |
| 38 | 3103 | 100.0 | 651 | 7 ADD06597 | Human Ckb |
| 39 | 3103 | 100.0 | 652 | 7 ADD06596 | Human Ckb |
| 40 | 3103 | 100.0 | 653 | 7 ADD06595 | Human Ckb |
| 41 | 3103 | 100.0 | 656 | 7 ADD06594 | Human Ckb |
| 42 | 3103 | 100.0 | 659 | 7 ADC16791 | Chimeric  |
| 43 | 3103 | 100.0 | 659 | 7 ADC16793 | Chimeric  |
| 44 | 3103 | 100.0 | 662 | 7 ADC16789 | Chimeric  |
| 45 | 3103 | 100.0 | 662 | 7 ADC16787 | Chimeric  |

ALIGNMENTS

RESULT 1  
AAP90388  
ID AAP90388 standard; protein; 585 AA.

XX AAP90388;  
AC  
XX 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 01-NOV-1989 (first entry)  
XX  
XX Mature human serum albumin polypeptide.  
XX Human serum albumin; mature protein; new polypeptides; plasma expanders.  
XX Homo sapiens; (Human).  
OS  
XX EP322094-A.  
FN  
XX 28-JUN-1989.  
PD  
XX 25-OCT-1988; 88EP-00310000.  
PF  
XX 30-OCT-1987; 87GB-00025529.  
PR  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA  
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
PI  
XX WPI; 1989-186464/26.  
DR  
XX N-ESDB; AAN90128.  
PT  
XX New N-terminal fragments of human serum albumin - esp. useful as blood  
PS plasma expanders.  
PT Disclosure; Fig 2; 20pp; English.

Mature protein of human serum albumin (see corresp. AAN90128). Used to make new N-terminal fragments which are used as plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

Query Match 100.0%; Score 3103; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.  
OX NCBI\_taxid=8508;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;  
RT "Partial mRNA sequence for tuatara A/B serum albumin";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF375973; AAM46106.1; -;  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR00264; Serum albumin.  
DR Pfam; PF00273; transport prot; 2.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN\_2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
FT NON TER 1  
FT SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;  
Query Match 33.7%; Score 1045; DB 13; Length 400;  
Best Local Similarity 47.7%; Pred. No. 2.8e-75;  
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;  
QY 196 QLKCAISLQKFGERAFKAWARLSQRPKAEFAEVSKLVTDLTKVHTCCGDLLECAD 255  
DB 3 EKHSQGLKSGFERAFQANKLVLSQKFPKAPPEIHKVLTATKLVKQECCHGDMIECLD 62  
QY 256 DRADLAKTICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPADLPDLAADFVESKDV 315  
DB 63 DRVENAYICSKQAVFSSKIKCCCKPIVDREVCIQLADLDEKPADLPFIAGQYIESTEV 122  
QY 316 CKNYAEAKDVFLGMLFYEARHPDYVSVLLRLAKTYETTLKCCAAADPHECYAKVFD 375  
DB 123 CKHYEGKDVFLAHFVYYSRHPFSSQMLRTGKGYQDILDKCKTENPPECYKAGE 182  
QY 376 EFKPLVERPQNLKONCELFEQLGEYKFNALLVRYTKKPOVSTPTLVEYSRNLGKVG 435  
DB 183 ELARHQESQELIKTHCSFTSQGKDPQOMVLVRYTKKPOLPAEBLIEISKLTGVGV 242  
QY 436 KCKKHPKAKRMPCAEDYLSVLNQLCVLHKTVPVSDRVTKCTESLVNRRPCFSALEVDE 495  
DB 243 KCCPLSEDKRLSSERKLSMWLFEICRQHEASPVNNHVTCCCTDSYSEMRPCFTKLGVD 302  
QY 496 TVVPKEFNAETTFHADICTLSEKROIKQTALVELVKHKPKATKQOLKAVMDDFAAV 555  
DB 303 SYVPEFCFSTFLFDEQLCTAPEEARLKKQLTLVLKLIQLKQIEDQKLVTDYHAME 362  
QY 556 EKCKKADDKETCFAEERGKLVAAASQAALGL 585  
DB 363 EKCCQAKENKQECFSTEGEKLQEGKALLGV 392

Search completed: April 19, 2004, 11:59:55  
Job time : 228.789 secs



```
QY 16 EENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADESAENCDKSLHTLFGDKLC 75
Db 40 EMNLVLDLATIFFAFVGHGATYETIRKMWANDVLAIVIKPTDGEHPTGCLNQMSAFLEEC 99
QY 76 TVATLRITYGEMADCCAKQPERNECFLOHKDNDPNLRLVRPEVDVMTAFDNEETFL 135
Db 100 HEKEIFKYG-LSOCCSRSEERHSCIPAHKKAAPALPPFOAPEPATSCKAYEEDRETFM 158
QY 136 KKYLYEIAARRHPYFAPELIFFAKRYKAAFTTECCQADKACALLPKLDELDRGKASSAK 195
Db 159 NKCIYEISRRHPFLYAPILISLAAGYDKMPLCCKAENAAECFQTKTASITKELKSSLL 218
QY 196 ORLKASLOKGERAFKAWARLVSQRPFKAEFAEVSUKLVTDLTKVHTECHGDLLECAD 255
Db 219 NOHVCVKRTPEPLIFQAITVAKLSQKPPRANFTBIOKLVDVAHIHEECRGNVLECLQ 278
QY 256 DRADLAKYICENQDSISSKLECECKPPLLEKSHCIAEVNDEMADLPSLAADPVEKOV 315
Db 279 DGEKIMSYICRGALSSKIAECCCKPLTELGHCIIHAENDDKPEGLSPTLNRLPLGRDF 338
QY 316 CKNYAEAKDVLGMFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYAKVFD 375
Db 339 NOFSAENIIFLASVHYSRRHPELAVPILRVKGVQEVLEKCAOSEKPLEQDKEE 398
QY 376 EFKPLVEEPQNLKONCELFEQLEKQYKFNALLVRYTKVPQVSTPTLVEVSRNLGKVG 435
Db 399 ELQKTYQSSQAMAKSCGLFQKLGQYVLQNAFLVSVYKVPOLTSAEALMALTEKMAATAA 458
QY 436 KCKHPEAKRMPCADYLSVNLQNLVLEKTPVSDRYTKCTESLVNRRPCSALEVDE 495
Db 459 TCCQLSEDKQLACGGAADLIIGHLCIRHEVNPVNPVGVHGCNSSYANRRPCFSLSLVIDE 518
QY 496 TVVPKEFNAETFTFHADICTLSEKEROIKKQATLVLELVKHPKATKEQLKAVMDFAFV 555
Db 519 TVVPSFSADKIFPKDLQCOGVALQTKQBFNLVNLKQKFIETEQLLEAVADFSGLL 578
QY 556 EKCKKADKETCFABEGKLVAAQALGL 585
Db 579 ETCCQEQEQEVCFABEGPKLISKTRDALGV 608

RESULT 12
ID Q8UW06 PRELIMINARY; PRT; 624 AA.
AC Q8UW06;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OC NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
```

```
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08933BF4953EF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Fred.No.2.1e-78; Indels 12; Gaps 5;
Matches 214; Conservative 102; Mismatches 240;

QY 14 LGEENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADESAENCDKSLHTLFGDK 73
Db 45 IGVEHAKALAMALFQMLSKPHHEQVQVRVNVMDIADLCGRGAKHGDCGKSVMTIILNE 104
QY 74 LCTVATLRITYGEMADCCAKQPERNECFLOHKDNDP-NLRLVRPEVDVMTAFDNEE 132
Db 105 ICKTPEPENPEKYPFHEGCKCKEDPERHKCFIEHKSPTDKERTYEVKPSPEQICKOHAEND 164
QY 133 TFLKLYLYEIAARRHPYFAPELIFFAKRYKAAFTTECCQ-AADKAAACLLPKLDELDEGKA 191
Db 165 EFLGHYIHKVASHHTMYPPALISLTLHPDGIHVSHCCXDEATVGCQLSEKMPAHKEEVEH 224
QY 192 SSAKQRLKASLOKGERAFKAWARLVSQRPFKAEFAEVSUKLVTDLTKVHTECHGDL 251
Db 225 VCAVQKHNCYIILQNFNERALRASKAHACSKPFASFENVQRLTDGIHVLHQTCCGGNM 284
QY 252 ECADDRADLAKYICENQDSISSKLECECKPPLLEKSHCIAEVNDEMADLPSLAADPVE 311
Db 285 ACMAERKLUUTQTCCK-----KCCCKPVLERSECIVRLPNDKPADLSPEVRYFD 336
QY 312 SKDVCNRYAEAKDVLGMFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db 337 DPEVCKRFXEEGDAFWGRFLCDYAKIHPHSAELNLRISAGLEKAYKTCCEGAHNECIA 396
QY 372 KVDFEPKPLVEEPQNLKONCELFEQLEKQYKFNALLVRYTKVPQVSTPTLVEVSRNLG 431
Db 397 KEBETLRHEFEASKTKLTKTCGALEKLGYPHFONIMIVRYTGILPRSSDAFLIYIKTLT 456
QY 432 KVGSKCKCKHPEAKRMPCADYLSVNLQNLVLEKTPV-SDRYTKCTESLVNRRPCFSA 490
Db 457 NIGQCKCLFEDQMPCEGGLQWVFAQIC-QNQTFFENKLAHCKDSLSFTTFCFAA 515
QY 491 LEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKQATLVLELVKHPKATKEQLKAVMD 550
Db 516 LTVDYTVVPAPVTAESFNFNDEFTSEADLQAKOTFLHLVTRHPKTIDEOVKTISEK 575
QY 551 FAFVEKCKKADKETCFABEGKLVAA 578
Db 576 FLAMCGGCKQADQNECFATEGAKLVEA 603

RESULT 13
ID Q8BK65 PRELIMINARY; PRT; 605 AA.
AC Q8BK65;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alpha-fetoprotein precursor.
GN AFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK076053; BAC36150.1; -.
DR MGD; MGI:87951; Afp.
DR GO; GO:0005615; C:extracellular space; IDA.
```

```
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;
Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 8.5e-91;
Matches 237; Conservative 110; Mismatches 238; Indels 6; Gaps 3;

QY 1 DAHKSEVAHRFKDL-----GEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVAD 56
DB 28 EGHVDNPPHLLIGDLIPMIGVDNSKGLVLAASQMLPLCFYSEHLQVRVEDVMQIADLCAG 87
QY 57 ESAENCDSKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDN-PNLPRL 115
DB 88 ARHANCASPMIILDELCKPENAEKYPFHQSCCKEDPEHKCFVEHKVANHEELTKY 147
QY 116 VRPEVDMCTAFHNDSETFLKYLVEIARRHYPFYFAPELLPFAKRYKAAFTCCOAAKA 175
DB 148 VRPAPEQICKDHAENRGPLIARYIFMLAIGHFMYTIPAILGPAQDFDGIIVSHCCXDVETA 207
QY 176 A-CLLPKLDLDEGRKASSAKQRLKCAQKQGERAFKAWAVARLSQRPFAEFAEVSKL 234
DB 208 GQCFNDKMPHEHQVEYVICALQKNCYIILQDPKRALTAAYKAVQASQKFPFLASFENVQII 267
QY 235 VTDLTKVHTCCCHGDLLECADRADLAKYICNQDSISSKLECKECPKPLEKSHCIAEVE 294
DB 268 VPDTVHLQTCGCGDMWACMLERMKLTAKICEKQDELATHLKECCDKPILERSACIIRLP 327
QY 295 NDEMPADLSLAADFVESKDVCNVAEAKDVLGMLFLEYARRHPDYSVVLRLRAKTYE 354
DB 328 NQKPADLSPKPHYIDDFEVCKLYTEGGDTMGRFLYECARRHQDYSPEMLRMGSGYE 387
QY 355 TTLEKCCAAADPHECVAKVDFEKPILVEBPQNLIKONCELPQGLGEYKFNQALLVRYTKK 414
DB 388 EFLKCCAAAGHNECLAETESLKEIESVTLTKNCGALDKLSYLFQNLILPKYVAR 447
QY 415 VPOVSTPFLVEVRNLGKSGCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRYT 474
DB 448 MPALSSQSLRIITIGECCHRPEDQOCSEGGIGVFGQICMKQKITPVNKA 507
QY 475 KCCTESLVNRRPCFSALEVDVETVPKEFNAETFFHADICTISEKERQIKKQATLVELVK 534
DB 508 QCCSHSLSSQTPCFSAIPVDETVPPLSVASFNFDELCTTSEPEQSKQKQVFLRLMK 567
QY 535 HKPKATYEQKAVMDPFAARVEKCKADKTCFAEKGKLVAAASQALGL 585
DB 568 QYPHMTDEQLKTCVNFVPMVDQCKADNHNCFALGAKLIDACKAILAV 618

RESULT 10
Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Nonneman D., Vallet J.L., Christenson R.K.;
RT "Mapping of the porcine alpha-fetoprotein (APP) gene to SSC8.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517770; AA066710.1; -
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN, 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;
Query Match 39.3%; Score 1218.5; DB 6; Length 610;
Best Local Similarity 39.7%; Pred. No. 6.2e-89;
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;

QY 16 EENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADSEANCDKSLHTLFGDKLC 75
DB 40 EMLVLDLTIFFAQFVQEAITYKEVNMVQVLDLTVIEKSTGSEQPACLENQVSVFLEEC 99
QY 76 TVATLRETYGEMADCCAKOEPERNECFLOHKDDN-PNLPRLVRPEVDMCTAFHNDSETFL 134
DB 100 HEBEIPKYG-LSHCCSQSGEERHNCFLARKKAAPASIPFPQVPEPVTSCAYEENRELF 158
QY 135 LKYLVEIARRHYPFYFAPELLPFAKRYKAAFTCCOAAKAACLLPKLDELDRDEGKASSA 194
DB 159 MTRYIYEIARRHYPFYFAPELLPFAKRYKAAFTCCOAAKAACLLPKLDELDRDEGKASSA 218
QY 195 KQRLKCAQKQGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTKVHTCCCHGDLLECA 254
DB 219 LNQHMCVTVMQFGARTFRAITVTKLSQKPPKANFTBIQKLVLDVAHIHECCRGNVLECL 278
QY 255 DRRADLAKYICNQDSISSKLECKECP-LLEKSHGICAEVENDEMPADLPSLAADFVESK 313
DB 279 QDAERVSVYVCSQDITLSSKIAECCKLPTTLEGGCIIHAENDKKEGSLSPNLRFLGER 338
QY 314 DVCQNVAEAKDVLGMLFLEYARRHPDYSVVLRLRAKTYETTTLEKCCAAADPHECVAKV 373
DB 339 DFNQLSREKDLSMARTYYSRRHPKLAVPVILRVAKGYOELLEKCSQSENPLECQDKG 398
QY 374 DFEKPLVBPQNLIKONCELPQGLGEYKFNQALLVRYTKKQVSTPFLVEVRNLGKV 433
DB 399 EEELEKVIQBSQALAKESGCLFQKLGLEYLQNAFLVAYTKAPQLTPPELMALTRKWT 458
QY 434 GSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRYTKCTESLVNRRPCFSALEV 493
DB 459 GAACHLSEBQLACGGAADLIIGQLCIHHEMPINPVGQCTSSYANRRPCFSSLV 518
QY 494 DETVPKEFNAETFFHADICTISEKERQIKKQATLVELVKHFKPKATKEQLKAVMDPFAA 553
DB 519 DETVVPFPFSDKIFPHKDLCAQAGVALQTMKQOFLINLVKQKPFQITEQLEAVIADFSG 578
QY 554 FVEKCKADKTCFAEKGKLVAAASQALGL 585
DB 579 LLEKCCQGEQEVCFAGEGPALEKTRASLG 610

RESULT 11
Q7TSF3 PRELIMINARY; PRT; 608 AA.
AC Q7TSF3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alpha-fetoprotein.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Scuriinae;
OC Marmota.
OX NCBI_TaxID=9995;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin H.L., Chen P.J., Wu H.L., Lee R.S., Chen D.S.;
RT "Molecular Cloning and Expression of Woodchuck Alpha-Fetoprotein.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034049; AAK55757.1; -
SQ SEQUENCE 608 AA; 67962 MW; EB46D7A37382B222 CRC64;
Query Match 38.1%; Score 1183.5; DB 11; Length 608;
Best Local Similarity 39.3%; Pred. No. 3.9e-86;
Matches 224; Conservative 108; Mismatches 237; Indels 1; Gaps 1;
```

```

Db      5  CLKSLDTITFLDEICHEEGFAKY-DLAACCAKAEVERKECLLAHKNATPOGIFAFQPGI 63
QY      121 DVMCTAFADNDETFKLYIARRHPYFAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
Db      64  EVSCKLYQDDRLTLGNVYIEVARHPVLPVPFATASLYDEALKTCCTQADKATCFHP 123
QY      181 KLDELDEGKASSAKQRLKCSLQKFGPRAPKAVARLSQRPFAEFAVSKLVTLTK 240
Db      124 RIPPLIEYKMSNGIQENTCGILKFGERTLKATKLVMSQKFPKADPATINKLVIDITH 183
QY      241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPA 300
Db      184 MHTCECRGDTLECLDRRLALTYTCSHKAOLSSKLPCTCEKSVLGERGECIVRLENDKPA 243
QY      301 DLPISLAADFVSKOVCKNYABAKVFLGMFLYIEYARRHPDYSVLLRLAKTYETTLK 360
Db      244 DLSERIAEVIDPHVCDHLAKEQDAFLAKFLYYSRRHPPELSTQILLGVGKYQELLERC 303
QY      361 CAADAPHECYAKVDFEFKPLVEEPONLJKONCELFQELGEYKFNALLVRYTKVPOVST 420
Db      304 KXTDNPPECYGQAEADLKKHQAQFOLVQNCDDLNTLGGVLFHALLIRTKRMPQITS 363
QY      421 PTLVEVSRNLGVSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
Db      364 BELIFVTR-ITKAASRCCEVSDVKLPCTEGVDFVLGQICORHORSSINWVQCQCSNS 422
QY      481 LVNRPCPSALEVDITYPKFENATFTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Db      423 YALRSLCITSLSGGDKFVPDIEFSADLFTFHEDLCHAAQDKQERRQQMIVNLVKKHNIT 482
QY      541 KQLKAVMDDDFAAFVEKCKADDKTCFAEKGKLVAAASQAAL 583
Db      483 KQLQTVFGFTKMTKECKCAEDHEACFEGEGLVAESSQAL 525

RESULT 8
Q8MJU5 PRELIMINARY; PRT; 609 AA.
AC Q8MJU5;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Alpha-fetoprotein.
GN AFP.
OS Canis familiaris (Dog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furuichi M., Neo S., Hiasue M., Tsuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.;
RT "Canine alpha-fetoprotein cDNA.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A5089789; BAC07513.1; -.
DR GO; GO:0003615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250CSAF2AF0 CRC64;

Query Match 40.0%; Score 1242; DB 6; Length 609;
Best Local Similarity 40.5%; Pred. No. 8.2e-91;
Matches 231; Conservative 113; Mismatches 225; Indels 2; Gaps 2;

QY      16 BENFKALVLIAPAYTLQCCPFEDHVKLVNEVTEFAKTCVADSSAENCDSKSLHTLFGDKL 75

```

```

Db      40  ENMLVDLATIPFAQFQVQATYKEVSMVKDILITVIEKSTGSEQPGGCLNQLPAFLKEIC 99
QY      76  TVATLRETYGEMADCAKQEPERNECFTLQKD-DNPNLPRLVRPEVDVWVCTAFHNEETVF 134
Db      100 HEKEISEKYG-LADCCSQSEERHNCFLAHKAAFPSPFPQVABEFTVTSCKAYEENRDMF 158
QY      135 LKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAADKAAACLLPKLDELDEGKASSA 194
Db      159 MNRYYIEIARRHPFLYAPTILSLAAHYGKIIPLCCKAENAVNECFQTKTSLITKELESRL 218
QY      195 KQRLKCSLQKFGPRAPKAVARLSQRPFAEFAVSKLVTLTKVHTTECHGDLLECA 254
Db      219 LNQHICAVMNEFGPRTFRITVTKLQSKFSKANFTIEIKLVLDVAHIHECCRGVLECL 278
QY      255 DBRADLAKYICENODSISSKLKECEKPLEKSHCIAEVENDEMPADLPSLADFVESKD 314
Db      279 QDGEKIMSYICQQDITLSKIDCCCKLPILLEGQCIHAENDGKPEGLSPNLRLEED 338
QY      315 VCKNVAEADVFLGMFLYIEYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYAKVF 374
Db      339 FNOFSSREKDLFWARTYYSRRHTKLAVPVLRVAKGYOELLEKCSQSENPLECQDKGE 398
QY      375 DEFKPLVEEPONLJKONCELFQELGEYKFNALLVRYTKVQVSTPTLVEVSRNLGVK 434
Db      399 EELEKYIQESQALAKESCGLFQKLGSEYLLQNAFLVAYTKKAPQLTPPELMATRKMATAA 458
QY      435 SKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTESLVNRRPCFSALEYD 494
Db      459 ATCCQLSEDRQLACGGAADLIIGQLCIRHEETPINPGVQCSCSSYANRRPCFSSLVVD 518
QY      495 ETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGKPKATKEQLKAVMDDDFAAF 554
Db      519 ETYPSPFSADKEIFHKDLCOAGVALQTMKQOFLINLVKQKPOITEEQLEAVIADFSG 578
QY      555 VEKCKADDKTCFAEKGKLVAAASQAALGL 585
Db      579 LEKCCQGEQAECEEGPKLISKTRALGV 609

RESULT 9
Q8UW05 PRELIMINARY; PRT; 626 AA.
AC Q8UW05;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma maculatum (Spotted salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;
OC Ambystoma.
OX NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazarteeren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217183; AAL56646.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SIGNAL.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 626 SERUM ALBUMIN.

```



```

DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 1687648333EFAE8D CRC64;

Query Match 64.2%; Score 1991; DB 4; Length 417;
Best Local Similarity 95.2%; Pred. No. 6.5e-151;
Matches 381; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

QY 186 RDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPKAEFAVSKLVTDLTQVHTEC 245
DB 24 RDAKSEVAHR-----FKDLGSENFKAWAVARLSQRFPKAEFAVSKLVTDLTQVHTEC 77
QY 246 CHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSL 305
DB 78 CHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSL 137
QY 306 AADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKCAAAD 365
DB 138 AADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKCAAAD 197
QY 366 PHECYAKVDFEKLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVSTPTLVE 425
DB 198 PHECYAKVDFEKLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVSTPTLVE 257
QY 426 VSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRR 485
DB 258 VSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRR 317
QY 486 PCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQFALVELVKHKPKATKEOLK 545
DB 318 PCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQFALVELVKHKPKATKEOLK 377
QY 546 AVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 378 AVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 417

RESULT 6
Q8IU7 PRELIMINARY; PRT; 396 AA.
AC Q8IU7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005815; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 60.1%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.6%; Pred. No. 6.7e-141;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGSENFKALVLIAPQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 84

us-09-832-929-18.rspt
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKYLIEYARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 180
DB 145 DVNCTAFHDNEETFLKYLIEYARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 163
QY 181 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPKAEFAVSKLVTDLTQ 240
DB 164 -----FKDLGSENFKAWAVARLSQRFPKAEFAVSKLVTDLTQVHTEC 163
QY 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 164 -----FKDLGSENFKAWAVARLSQRFPKAEFAVSKLVTDLTQVHTEC 163
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK 360
DB 164 -----FKDLGSENFKAWAVARLSQRFPKAEFAVSKLVTDLTQVHTEC 171
QY 361 CAADPHECYAKVDFEKLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 172 CAADPHECYAKVDFEKLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 231
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 291
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQFALVELVKHKPKAT 540
DB 292 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQFALVELVKHKPKAT 351
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 352 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 396

RESULT 7
Q8JIA9 PRELIMINARY; PRT; 527 AA.
AC Q8JIA9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J.; Brennan S.O.; George P.M.; Chambers G.K.;
RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375971; AA046104.1; -
DR GO; GO:0005815; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 3.6e-95;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

```

```

DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER 1
SQ SEQUENCE 576 AA; 65002 MW; F857333E99AE37F04 CRC64;

Query Match 75.3%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred. No. 2.5e-178;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

Qy 9 HRFKDLGEENFKALVLIAFAYQLQQCFPEFDHVLKLVNEVTEFAKTCVADESAENCDSLSHT 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 NRYNDLGEQHFQGLVLIAFQYQLKQCSYDEHAKLVQEVTDFAKTCVADESAANCDSLSHT 60

Qy 69 LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEVDVMCTAFH 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LFGDKLCAIENLRENYGELADCTKQEPERNECFLOHKDDNPNLPLRPERPEAEANCTSFK 120

Qy 129 DNBETFLKKYLYIARIHPHYFYAPPELLFFAKRYKAAFTCCQAADKAACLLPKLDELRLDE 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 ENPTTFMGHYLHEVARRHPHYFYAPPELLYYAEQYNEILITQCCAEADKESCLTPKLDGVKEK 184

Qy 189 GKASSAKQRLKASLQKFGGRAPKANAVARLSORFPKAEFAEVSUKLVTDITKVTETCHG 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 ALVSSVRQRMKCSMQKFGGRAPKANAVARLSQTFPNADFAETIKLATDITKVNKECCHG 240

Qy 249 DLLECADDRADLAKYICENQDSISSLKECCCEKPLLEKSHCTAEVNDENPADLPSLAAD 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 DLLECADDRADLAKYMCENQATISSLQTCDFELLKKAHCLSEVEHDTWPADLPATAAD 300

Qy 309 FVSKQVCKNYBAKDVFLGMYEYARRHPDYSVLLLRLLAKTYETTLPEKCCAAADPHE 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 FVEDQECVCKNYBAKDVFLGMYEYARRHPDYSVLLLRLLAKTYEATLEKCCAEANPPA 360

Qy 369 CYAKVDFEFPKLVPEEPONLTKONCELPEQLGEYKFNQALLVRYTKYPOUSTPTLVESR 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 CYGTVLAEFQPLVEEPKNLVKTCDLYEKLGEGYFQNALIVRYTKQAPQVSTPTLVEAAR 420

Qy 429 NLGKVGSKCCKHPEAKRMPCAEDYSVLVNLQCVLHKEKTPVSDRVTCKCTESLVNRRPCF 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 NLGRVGTCKCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTCKCSGSLVERRPCF 480

Qy 489 SALEVDETYVPKFNPAETPTFHADICTLSKEKQIKKOTALVELVKKHKKATKQOLKAVM 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 SALTVDITYVPKGFKAETFTFHSDICTLPEKEKQIKKOTALABELVKKHKKATABQLKAVM 540

Qy 549 DDFAAFVEKCKCKADDKETCFABEGCKLVAASQAAL 583
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 DDFAQFLDTCCKAADKDTCTESTGPNLVTRCKDAL 575

RESULT 5
Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport prot; 2.
DR PRINTS; PR00402; SERUMALBMIN

```

```

121 DAVCTAFHNEETFLKYLIEIARRHPYFAPPELLYFAEYRGVTECEADKAACLTIP 180
181 KLDELDEKASSAKORLKCASLOKFGGRAPKAWAVARLSORFPKABEAYSKLVTDLT 240
181 KVALREKVLASSAKERLKCASLOKFGGRAPKAWAVARLSORFPKADPAEISKLVDLAK 240
241 VHTCECHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAFVENDMPA 300
241 IHKECHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAFVENDMPA 300
301 DPLSLAADFVESKDVCKNYABAKDVLGMLFVEYARRHPDYSVLLLRKATYETTLK 360
301 DPLSLAADFVESKDVCKNYABAKDVLGMLFVEYARRHPDYSVLLLRKATYETTLK 360
361 CAAADPHCEYAKVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
361 CATDDPPACYAHVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKAT 540
481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKAT 540
541 KEQIKAVMDDFAAVFEKCCADKCTCPABEGKLVAAQAAL 583
541 EBQLKTVMGDFGSDVFDKCCAREADKACFAEFGKLVAAQAAL 583

```

## RESULT 2

```

Q95VB7 PRELIMINARY; PRT; 608 AA.
ID Q95VB7
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staderker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AL008579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66B54 CRC64;

```

```

Query Match 80.7%; Score 2504; DB 5; Length 608;
Best Local Similarity 76.3%; Pred. No. 9.9e-192;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGBENFKALYLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
25 DAHKEVAHRFKDLGBENFKALYLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 84
61 NCDKSLHTLFGDKLCITVATREYVGMADCCAKQEPERNECTLOHNDPNLPRLVREPV 120
85 NCDKSLHTLFGDKLCITVATREYVGMADCCAKQEPERNECTLOHNDPNLPRLVREPV 144

```

```

QY 121 DVMTAFHNEETFLKYLIEIARRHPYFAPPELLYFAEYRGVTECEADKAACLTIP 180
DB 145 EAMCTSFQENAVTFMGHYLHEVARRHPYFAPPELLYFAEYRGVTECEADKAACITP 204
QY 181 KLDELDEKASSAKORLKCASLOKFGGRAPKAWAVARLSORFPKABEAYSKLVTDLT 240
DB 205 KIDALKEKALASSVNRKLCSSLQRFQGRAPKAWAVARLSORFPKADPAEITKLAIDL 264
QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAFVENDMPA 300
DB 265 LTBECCHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAFVENDMPA 324
QY 301 DPLSLAADFVESKDVCKNYABAKDVLGMLFVEYARRHPDYSVLLLRKATYETTLK 360
DB 325 DPLSLAADFVESKDVCKNYABAKDVLGMLFVEYARRHPDYSVLLLRKATYETTLK 384
QY 361 CAAADPHCEYAKVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
DB 385 CAEADPSACYKGVLDLDFQPLVEEPKLVKANCELFELKGEYGFQNALVRYTKVPQVST 444
QY 421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVRNLGKVGSKCCCHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKAT 540
DB 505 VVERPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKAT 564
QY 541 KEQIKAVMDDFAAVFEKCCADKCTCPABEGKLVAAQAAL 583
DB 565 GPQLRTVLGFEFTAFDLKCKCAEDKACFSDGSPKLVASSQAAL 607

RESULT 3
Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3
AC Q8C7H3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600BED3A61B4 CRC64;

```

```

Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.5e-181;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 225.789 Seconds  
(without alignments)  
817.479 Million cell updates/sec

Title: US-09-832-929-18  
Perfect score: 3103  
Sequence: 1 DAHSEVAHFKDLGEENFK.....TCFAEGKLVLAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | ID     | Description        |
|------------|--------|-------------|--------|--------|--------------------|
| 1          | 2614   | 84.2        | 584    | Q7YSG3 | Q7YSG3 felis silve |
| 2          | 2504   | 80.7        | 608    | Q95V57 | Q95V57 schistosoma |
| 3          | 2374   | 76.5        | 608    | Q8C7H3 | Q8C7H3 mus musculu |
| 4          | 2336   | 75.3        | 576    | Q8C7C7 | Q8C7C7 mus musculu |
| 5          | 1991   | 64.2        | 417    | Q86YG0 | Q86YG0 homo sapien |
| 6          | 1865.5 | 60.1        | 396    | Q81UK7 | Q81UK7 homo sapien |
| 7          | 1295.5 | 41.7        | 527    | Q81IA9 | Q81IA9 sphenodon p |
| 8          | 1242   | 40.0        | 609    | Q8MJU5 | Q8MJU5 canis famil |
| 9          | 1242   | 40.0        | 626    | Q8UW05 | Q8UW05 ambystoma m |
| 10         | 1218.5 | 39.3        | 610    | Q8MJ76 | Q8MJ76 sus scrofa  |
| 11         | 1183.5 | 38.1        | 608    | Q7TSF3 | Q7TSF3 marmota mon |
| 12         | 1087   | 35.0        | 624    | Q8UW06 | Q8UW06 ambystoma t |
| 13         | 1082   | 34.9        | 605    | Q8BK65 | Q8BK65 mus musculu |
| 14         | 1076   | 34.7        | 605    | Q8BK56 | Q8BK56 mus musculu |
| 15         | 1045   | 33.7        | 400    | Q81IA7 | Q81IA7 sphenodon p |
| 16         | 955    | 30.8        | 603    | Q9YGH6 | Q9YGH6 rana shqipe |

|    |       |      |      |    |        |                    |
|----|-------|------|------|----|--------|--------------------|
| 17 | 928.5 | 29.9 | 614  | 13 | Q91134 | Q91134 naja naja   |
| 18 | 888   | 28.6 | 406  | 13 | Q81IA8 | Q81IA8 hoplodactyl |
| 19 | 759   | 23.8 | 205  | 11 | Q8CG74 | Q8CG74 mus musculu |
| 20 | 713   | 23.0 | 417  | 11 | Q8R0J9 | Q8R0J9 mus musculu |
| 21 | 376.5 | 12.1 | 484  | 13 | Q9W6F5 | Q9W6F5 gallus gall |
| 22 | 372   | 12.0 | 491  | 11 | Q7TS97 | Q7TS97 mus musculu |
| 23 | 365   | 11.8 | 551  | 13 | Q42279 | Q42279 petromyzon  |
| 24 | 331   | 10.7 | 122  | 13 | Q90WZ8 | Q90WZ8 larus argen |
| 25 | 290   | 9.3  | 123  | 13 | Q90WZ6 | Q90WZ6 poephila gu |
| 26 | 284   | 8.5  | 135  | 11 | Q63205 | Q63205 rattus norv |
| 27 | 196   | 6.3  | 807  | 2  | Q83ZS7 | Q83ZS7 helicobacte |
| 28 | 188   | 6.1  | 1723 | 2  | Q9JMX8 | Q9JMX8 helicobacte |
| 29 | 184   | 5.9  | 1819 | 16 | Q9ZLVO | Q9ZLVO helicobacte |
| 30 | 184   | 5.9  | 1927 | 16 | Q25262 | Q25262 helicobacte |
| 31 | 162.5 | 5.2  | 44   | 6  | Q95WC2 | Q95WC2 equus cabal |
| 32 | 162.5 | 5.2  | 680  | 5  | Q9V6S8 | Q9V6S8 drosophila  |
| 33 | 161   | 5.2  | 1079 | 3  | Q96V11 | Q96V11 pneumocysti |
| 34 | 156   | 5.0  | 1026 | 3  | Q74669 | Q74669 pneumocysti |
| 35 | 154.5 | 5.0  | 3843 | 5  | Q9UD00 | Q9UD00 drosophila  |
| 36 | 153.5 | 4.9  | 661  | 5  | Q8MS79 | Q8MS79 drosophila  |
| 37 | 153.5 | 4.9  | 3843 | 5  | Q9VU94 | Q9VU94 drosophila  |
| 38 | 153   | 4.9  | 62   | 6  | Q8MIL1 | Q8MIL1 sus scrofa  |
| 39 | 150   | 4.8  | 1065 | 3  | Q01828 | Q01828 pneumocysti |
| 40 | 149   | 4.8  | 40   | 6  | Q9TRA5 | Q9TRA5 oryctolagus |
| 41 | 148   | 4.8  | 1028 | 3  | Q74668 | Q74668 pneumocysti |
| 42 | 142.5 | 4.6  | 2756 | 10 | Q9LJ60 | Q9LJ60 arabidopsis |
| 43 | 141.5 | 4.6  | 1560 | 5  | Q26644 | Q26644 strongyloce |
| 44 | 138.5 | 4.5  | 1069 | 3  | Q96V12 | Q96V12 pneumocysti |
| 45 | 137.5 | 4.4  | 1348 | 16 | Q8YK55 | Q8YK55 anabaena sp |

## ALIGNMENTS

### RESULT 1

Q7YSG3 PRELIMINARY; PRT; 584 AA.

ID Q7YSG3

AC Q7YSG3

DT 01-OCT-2003 (TREMREL. 25, Created)

DT 01-OCT-2003 (TREMREL. 25, Last sequence update)

DT 01-OCT-2003 (TREMREL. 25, Last annotation update)

DE Albumin (Fragment).

GN ALB.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI\_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,

RA Rumpold H., Valenta R., Spitzauer S.,

RT "Escherichia coli expression and purification of recombinant cat

RT albumin: IgE recognition, induction of basophil activation and

RT lymphoproliferative responses in atopic patients."

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ487677; CAD32275.1; -

FT NON TER

SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 84.2%; Score 2614; DB 6; Length 584;  
Best Local Similarity 81.6%; Pred. No. 1.5e-200;  
Matches 476; Conservative 54; Mismatches 53; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | DAHSEVAHFKDLGEENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADESAAE | 60  |
| Db | 1   | EAHGEFAHFNLDGEEHFGVLVAESQVLQCCPDHVKLVNEVTEFANGCVADOSAA    | 60  |
| QY | 61  | NCDKSLHTFGDKLCTVATIRETYGENADCCAKOBERNECFLOHKDDNNLRLVPEV   | 120 |
| Db | 61  | NCDKSLHTFGDKLCTVATIRETYGENADCCAKOBERNECFLOHKDDNNLRLVPEV   | 120 |
| QY | 121 | DVMTAFDNEETLKKVLYEIARRHFFYFAPELLFAKRYKAAFTCCCAADCAACLLP   | 180 |

```

FT DISULFID 302 313
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CARBOHYD 251 251
FT
FT N-LINKED (GLCNAC. . .).
FT /FTid=CAR_000070.
FT A -> G (in dbSNP:1057173).
FT /FTid=VAR_012049.
SQ SEQUENCE 609 AA, 68677 MW, 4945820E1C2D4F CRC64;

Query Match 40.3%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.9%; Pred. No. 5.5e-73;
Matches 235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;

QY 3 HXSE-----VAHFRDGLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADE 57
DB 22 HNEVGLASILDSYQCTAEISLADLATIFFAQFVEATYKEVSKVAKDALTAEKPTGDE 81
QY 58 SAENCDKSLHTLFGKLCVTATRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLV 116
DB 82 QSSGCLNQLPALBELCHEKEILEKYGH-SDCCSQSBEGRHNCFLAHKXTPASIPLFQ 140
QY 117 REVDVMCTAFHNDRETFKXLYLRIARHPYFAPPELLFAKRYKAFAFTCCQAADKAA 176
DB 141 VPEPVTSCEAYBEDRETFNKKFIYIARHPPLYPAPTILLWAARYDKIIPSCCKAENAVE 200
QY 177 CULPKLDELRDEGKASSAQRLKASLQKGERAFKAWARLSQRFPPKAEFAEVSKLVT 236
DB 201 CPTKAAATVTKELRESSLLNHACAVMKNFGTRTQAIVTKLSQKTKVNFTEIQKLV 260
QY 237 DLTKVTECHGDLLECADDRADLAKYICENQDSISSKKECCEKPLLEKSHCIAEVEND 296
DB 261 DVAHVHEHCRCGVDLCLQDGKIMSYICSQDITLSNKITECCKLTLLERGQCIIHAEND 320
QY 297 EMPADLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLILAKTYET 356
DB 321 EKPEGLSPNINRFLGDRDNQFSSGKKNIFLASFVHEYSRRHPQLAVSVILAVAKGYQEL 380
QY 357 LEKCCAAADPHECYAKVDFEPLVEEPQNLKQNCLEPEQLGEYKFNALLVRYTKVP 416
DB 381 LEKCFOTENPLECQDKGEBELQYICESQALAKRSQGLFQKLGEYLLQNAFLVAYTKAP 440
QY 417 QVSTPTLVEYSRNLGVSKCKKHPKAEADYLSVVLNQLCVLHEKTPVSDRVTKC 476
DB 441 QLTSSSELMAITRKGAATAATCCQLSEDKLLACGGAADIIIGHLCIRHEWTPVNFVGQC 500
QY 477 CTESLVNRRPCFSALEVDETYVPKFNPAETFTFHADICTLSEKEROIKQTALVELVKHK 536
DB 501 CTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCOAGVALQTMKQEFLLINLVKQK 560
QY 537 PKATKEQLKAVMDDDFAFVEKCKKADDKETCFAEKGKLVAAASQAALGL 585
DB 561 PQITEGLEAVIADFSGLLEKCCQGGQEQEVCFABEGQKLISKTRAAALGV 609

```

Search completed: April 19, 2004, 11:52:43  
Job time : 44.6731 secs

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
FAhey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Scherth A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 1-28 FROM N.A.  
RX MEDLINE=93279385; PubMed=7684942;  
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,  
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;  
RT "A G--A substitution in an HNF I binding site in the human alpha-  
fetoprotein gene is associated with hereditary persistence of alpha-  
fetoprotein (HPAFP).";  
RL Hum. Mol. Genet. 2:379-379(1993).  
RN [5]  
RP SEQUENCE OF 429-556 FROM N.A.  
RX MEDLINE=83158778; PubMed=6187626;  
RA Beattie W.G., Dugaiczky A.;  
RT "Structure and evolution of human alpha-fetoprotein deduced from  
partial sequence of cloned cDNA.";  
RL Gene 20:415-422(1982).  
RN [6]  
RP PARTIAL SEQUENCE OF 19-609.  
RX MEDLINE=91242409; PubMed=1709810;  
RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F.,  
RA Ceccarini C., Terrana B.;  
RT "Human alpha-fetoprotein primary structure: a mass spectrometric  
study.";  
RL Biochemistry 30:5061-5066(1991).  
RN [7]  
RP PRELIMINARY SEQUENCE OF 19-35.  
RX MEDLINE=77242506; PubMed=70228;  
RA Yachnin S., Heu R., Heinrikson R.L., Miller J.B.;  
RT "Studies on human alpha-fetoprotein. Isolation and characterization  
of monomeric and polymeric forms and amino-terminal sequence  
analysis.";  
RL Biochim. Biophys. Acta 493:418-428(1977).  
RN [8]  
RP PRELIMINARY SEQUENCE OF 19-38.  
RX MEDLINE=78001760; PubMed=71198;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Comparative chemical structures of human alpha-fetoproteins from  
fetal serum and from ascites fluid of a patient with hepatoma.";  
RL Cancer Res. 37:3663-3667(1977).  
RN [9]  
RP PRELIMINARY SEQUENCE OF 19-39.  
RX MEDLINE=75018719; PubMed=4138095;  
RA Ruoslahti E., Pitko H., Vaheri A., Seppala M., Virolainen M.,  
RA Kontinen A.;  
RT "Alpha fetoprotein: structure and expression in man and inbred mouse  
strains under normal conditions and liver injury.";  
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).  
RN [10]  
RP GENE STRUCTURE.  
RX MEDLINE=85182629; PubMed=2580830;  
RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,  
RA Tamaoki T.;  
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'  
flanking region.";  
RL J. Biol. Chem. 260:5055-5060(1985).  
RN [11]  
RP METAL-BINDING.  
RX MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
RL Cancer Res. 38:3483-3486(1978).  
RN [12]  
RP BILIRUBIN-BINDING.  
RX MEDLINE=80001710; PubMed=89900;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-  
binding ability.";  
RL Cancer Res. 39:3571-3574(1979).  
RN [13]  
RP SULFATION.  
RX MEDLINE=86042625; PubMed=2414772;  
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
RT "Tyrosine sulfation of proteins from the human hepatoma cell line  
HepG2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
RN [14]  
RP FUNCTION: Binds copper, nickel, and fatty acids as well as, and  
bilirubin less well than, serum albumin. Only a small percentage  
(less than 2%) of the human AFP shows estrogen-binding properties.  
CC - SUBUNIT: Dimeric and trimeric forms have been found in addition  
to the monomeric form.  
CC - SUBCELLULAR LOCATION: Secreted.  
CC - TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and  
yolk sac.  
CC - DEVELOPMENTAL STAGE: Occurs in the plasma of fetuses more than 4  
weeks old, reaches the highest levels during the 12th-16th week of  
gestation, and drops to trace amounts after birth. The serum level  
in adults is usually less than 40 ng/ml. Afp occurs also at high  
levels in the plasma and ascitic fluid of adults with hepatoma.  
CC - PTM: Independent studies suggest heterogeneity of the amino-  
terminal sequence of the mature protein and of the cleavage site  
of the signal sequence.  
CC - PTM: Sulfated.  
CC - SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC - SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; M10949; AAAS1674.1; -;  
EMBL; M10950; AAAS1675.1; -;  
EMBL; V01514; CAA24758.1; -;  
EMBL; M16110; AAB58754.1; -;  
EMBL; BC027881; AAH27881.1; -;  
EMBL; Z19532; CAA79592.1; -;  
PIR; A26624; FPHU.  
HSP; P02768; IE78.  
GlycoSuiteDB; P02771; -;  
Sienna-2DPAGE; P02771; -;  
Gene; HGNC:317; AFP.  
MIN; 104150; -;  
InterPro; IPR000264; Serum\_albumin.  
Pfam; PF00273; transport\_prot; 3.  
PRINTS; PR00802; SERUMALBUMIN.  
ProDom; PD002486; Serum\_albumin; 1.  
SMART; SM00103; ALBUMIN; 3.  
PROSITE; PS00212; ALBUMIN; 2.  
KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;  
KW Signal; Polymorphism.  
FT SIGNAL 1 18  
FT CHAIN 19 609 ALPHA-FETOPROTEIN.  
FT DOMAIN 20 205 ALBUMIN 1.  
FT DOMAIN 212 397 ALBUMIN 2.  
FT DOMAIN 404 595 ALBUMIN 3.  
FT METAL 22 22 COPPER AND NICKEL.  
FT DISULFID 99 114  
FT DISULFID 113 124  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303







[4] SEQUENCE OF 99-516 FROM N.A.  
RP MEDLINE=88216123; PubMed=2452956;  
RA Minghetti P.P., Law S.W., Dugaiczyk A.;  
RT "The rate of molecular evolution of alpha-fetoprotein approaches that  
of pseudogenes";  
RL Mol. Biol. Evol. 2:347-358(1985).  
RN [5]  
RN SEQUENCE OF 477-551 FROM N.A.  
RP STRAIN=BALB/c;  
RC MEDLINE=90269606; PubMed=1971802;  
RA Boccaccio C., Dechatrete J., Meunier-Rotival M.;  
RT "Empty and occupied insertion site of the truncated LINE-1 repeat  
located in the mouse serum albumin-encoding gene";  
RL Gene 88:181-186(1990).  
RN [6]  
RN SEQUENCE OF 25-44.  
RP TISSUE=Liver;  
RC MEDLINE=93162044; PubMed=1286668;  
RA Giometti C.S., Taylor J., Tollaksen S.L.;  
RT "Mouse liver protein database: a catalog of proteins detected by two-  
dimensional gel electrophoresis";  
RL Electrophoresis 13:970-991(1992).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ011413; CA09617.1; -;  
DR EMBL; AK010025; BAB26650.1; -;  
DR EMBL; BC049571; AAH49971.1; -;  
DR EMBL; M61111; AAA37190.1; -;  
DR EMBL; X13060; CAA31458.1; -;  
DR PIR; A05139; A05139.  
DR HSSP; P02768; 1E7B.  
DR SWISS-2DPAGE; P07724; MOUSE.  
DR MGD; MGI:87991; Alb1.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; transport\_p107\_3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608  
FT DOMAIN 25 205  
FT DOMAIN 212 397  
FT DOMAIN 404 595  
FT METAL 27 27  
FT METAL 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 340 385  
BY SIMILARITY.

FT DISULFID 384 393 BY SIMILARITY.  
FT DISULFID 416 462 BY SIMILARITY.  
FT DISULFID 461 472 BY SIMILARITY.  
FT DISULFID 485 501 BY SIMILARITY.  
FT DISULFID 500 511 BY SIMILARITY.  
FT DISULFID 538 583 BY SIMILARITY.  
FT DISULFID 582 591 BY SIMILARITY.  
FT CONFLICT 27 27 H -> D (IN REF. 6).  
FT CONFLICT 33 33 H -> D (IN REF. 6).  
FT CONFLICT 41 41 Q -> I (IN REF. 6).  
SQ SEQUENCE 608 AA; 68692 MW; 2927F7C7EED3A61B4 CRC64;  
  
Query Match 76.6%; Score 2378; DB 1; Length 608;  
Best Local Similarity 72.4%; Pred. No. 2.6e-145;  
Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;  
  
QY 1 DAKSEVAHRFKDLGEENFKALVLIAFAQYLOOCPEPDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 EAHKSEIAHRYNDLGEQHFGLVLIAFSOLQKSYDEHAKLVQEVTDFAKTCVADESAA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECFLOHKDDNPSPFPERPEA 144  
QY 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKACLP 180  
DB 145 EAMCTSPKENPTTFMGHYLHEVARRRHPYFYAPPELLYYAQQYNELITQCCAEADKESCLTP 204  
QY 181 KLDELKDEGKASAKORLKASLOKGERAFKAWAVARLSQREPKAEFAEVSKLVDLT 240  
DB 205 KLDGVKEKALVSSVRQMKSSMQKTFGERAFKAWAVARLSQTFPNADFAITKLADLT 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECECKPILLEKSHCIAEVENDEMPA 300  
DB 265 VNKECCGHDILLECCADRAELAKYCNENQAISLQTCCKPILKKAHCLSEVEHDTMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMPFLYEYARRHPDYSVLLILRLAKTYETTLK 360  
DB 325 DLPAIAADFVEDQVCKNYAEAKDVFLGTFLYEYARRHPDYSVLLILRLAKTYETTLK 384  
QY 361 CAADPHECAKVFDEPKPLVERPQNLKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420  
DB 385 CREANPPACVGTVLAERQPLVEEPKLVKTCNCDLYEKLGEYQNALVRYTKAPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRTVKCTES 480  
DB 445 PTLVEAARNLGRVGTCKCTLPDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTCKCSGS 504  
QY 481 LVNRPFCFSALEVDETVVPKEENAETTFHADICTLSEKERQIKQTALVELVKHPEKAT 540  
DB 505 LVERRPCFSAVTVDETVVPKEFKAETFFHSDICTLPEKEKQIKQTALAEVLVKHPEKAT 564  
QY 541 KEQLKAVMDDPAFAFVEKCCKADDKETCFABEGKLVAAASQAAL 583  
DB 565 AEQLKTVMDPFAQLFDTCCKAADKDTCFSTEGPNLVTRCKDAL 607  
  
RESULT 13  
ID ALBU CHICK STANDARD; PRT; 615 AA.  
AC P19121;  
DT 01-NOV-1990 (Rel. 15, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).  
OS ALB.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.

DR PIR; JC5838; JC5838.  
 DR HSP; P02768; 1E7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport p107; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 609  
 FT DOMAIN 25 206  
 FT DOMAIN 213 398  
 FT DOMAIN 405 596  
 FT METAL 28 87  
 FT DISULFID 78 87  
 FT DISULFID 100 116  
 FT DISULFID 115 126  
 FT DISULFID 149 184  
 FT DISULFID 193 202  
 FT DISULFID 225 271  
 FT DISULFID 270 278  
 FT DISULFID 290 304  
 FT DISULFID 303 314  
 FT DISULFID 341 366  
 FT DISULFID 365 394  
 FT DISULFID 417 463  
 FT DISULFID 462 473  
 FT DISULFID 486 502  
 FT DISULFID 501 512  
 FT DISULFID 539 584  
 FT DISULFID 583 592  
 FT SEQUENCE 609 AA; 68940 MW; 9CA5F97F67EFLA48 CRC64;  
 Query Match 76.9%; Score 2387; DB 1; Length 609;  
 Best Local Similarity 73.9%; Pred. No. 6.9e-146;  
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;  
 QY 2 AHKSEVAHPRKIDGKNGKALVLIAPAOYLQCPEDHVKLVNEVTEFAKTCVADESSEN 61  
 DB 27 AHKSEVAHPRKIDGKNGKALVLIAPAOYLQCPEDHVKLVNEVTEFAKTCVADESSEN 86  
 QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEVD 121  
 DB 87 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEVD 146  
 QY 122 VMCTAFHDNETFLKYLVEIARRHPVVEYAPPELLFFAKRYKAAFTCCCOAADAACLLPK 181  
 DB 147 AMCTAFQENAEAFMGHYLHEVARHPYFGPELLYLADKITAVLTCCADDDKAGCLTEK 206  
 QY 182 LDELRDGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTKV 241  
 DB 207 LDALKERKALYSAYRQLKCSMMKFGRAFKAFAVAVMSQTFPNADFAETKLTATDLTKV 266  
 QY 242 HTECHGDLLECADDRLADLAKYICENQDSISSKLKCEKPELEKSHCAEVENDEMPAD 301  
 DB 267 TQECCHGDLLECADDRLADLAKYICENQDSISSKLKCEKPELEKSHCAEVENDEMPAD 326  
 QY 302 LPSLAADFVSKQVCKNYAKADVFLGWFLEYEARHPDYSVVLLLELAKTYETTLKCC 361  
 DB 327 LPALTADFVEDKQVCKNYAKADVFLGWFLEYEARHPDYSVVLLLELAKTYETTLKCC 386  
 QY 362 AAADPHSCYAKVFDEPKLVEEPQNLKQCELPQGEQFQNALVRYTKKQVQVSTP 421  
 DB 387 AEADPHACYGVHVFDEPKLVEEPQNLKQCELPQGEQFQNALVRYTKKQVQVSTP 446  
 QY 422 TLVEVSRNLGVKSGCKGHPKAMPKCAEDYLSVVLNQLCVLHEKTEPVSQVTKCCSTESL 481  
 DB 447 TLVEAARSLGVTHCCALPEKRLPCVEDYLSAILNRVCLHEKTPVSQVTKCCSGSL 506  
 QY 482 VNRPCFSALVDETVVPKFKAEFTTFHANICTLPEKEKQWKOTALAEVLVKKHPQATE 566

DB 507 VERRPCFSALVDETVVPKFKAEFTTFHANICTLPEKEKQWKOTALAEVLVKKHPQATE 566  
 QY 542 EQLKAVMDFFAFAFVEKCKKADKCKTCEFAEKGKLVAAASQAL 583  
 DB 567 EQLKAVMDFFAFAFVEKCKKADKCKTCEFAEKGKLVAAASQAL 608  
 RESULT 12  
 ALBU\_MOUSE  
 ID ALBU\_MOUSE STANDARD; PRT; 608 AA.  
 AC P07724; Q61802;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALBU OR ALB1 OR ALB-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6J; TISSUE=Tongue;  
 RA MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Kozaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant I.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirbl L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Bono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Hayashizaki Y.,  
 RA Yawshav-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RL "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RC TISSUE=Liver;  
 RX MEDLINE=89016582; PubMed=3174440;  
 RA Baldwin G.S., Weinstock J.;  
 RT "Nucleotide sequence of porcine liver albumin.";  
 RL Nucleic Acids Res. 16:9045-9045(1988).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X12422; CAA30970.1; -;  
 DR EMBL; N36787; AAA30988.1; -;  
 DR PIR; S01382; ABPGS.  
 DR HSP; P02768; IEVH.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transprot prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN\_3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 KW METAL-BINDING; LIPID-BINDING; REPEAT; SIGNAL; COPPER.  
 FT NON\_TER 1  
 FT SIGNAL <1 16 BY SIMILARITY.  
 FT PROPEP 17 22 BY SIMILARITY.  
 FT CHAIN 23 605 SERUM ALBUMIN.  
 FT DOMAIN 23 202 ALBUMIN 1.  
 FT DOMAIN 209 394 ALBUMIN 2.  
 FT DOMAIN 401 582 ALBUMIN 3.  
 FT METAL 31 31 COPPER (BY SIMILARITY).  
 FT DISULFID 75 84 BY SIMILARITY.  
 FT DISULFID 97 113 BY SIMILARITY.  
 FT DISULFID 112 123 BY SIMILARITY.  
 FT DISULFID 145 190 BY SIMILARITY.  
 FT DISULFID 183 198 BY SIMILARITY.  
 FT DISULFID 221 267 BY SIMILARITY.  
 FT DISULFID 266 274 BY SIMILARITY.  
 FT DISULFID 286 300 BY SIMILARITY.  
 FT DISULFID 299 310 BY SIMILARITY.  
 FT DISULFID 337 382 BY SIMILARITY.  
 FT DISULFID 381 390 BY SIMILARITY.  
 FT DISULFID 413 459 BY SIMILARITY.  
 FT DISULFID 458 469 BY SIMILARITY.  
 FT DISULFID 482 498 BY SIMILARITY.  
 FT DISULFID 497 508 BY SIMILARITY.  
 FT DISULFID 535 580 BY SIMILARITY.  
 FT DISULFID 579 588 BY SIMILARITY.  
 FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).  
 FT SEQUENCE 605 AA; 69410 MW; 3E556BDD1A1FAFF CRC64;  
 Query Match 77.7%; Score 2411.5; DB 1; Length 605;  
 Best Local Similarity 76.0%; Pred. No. 1.9e-147;  
 Matches 436; Conservative 67; Mismatches 70; Indels 1; Gaps 1;  
 QY 1 DANKSEVAHFKDGLGEENFVALVIAFAVLQOCPEEDHVKLVNEVTEFAKTCVADSAE 60  
 DB 23 DTKSEIAHFKDGLGEYFGVLVIAFSLQOCPEVEHVKLVREVTEFAKTCVADSAE 82  
 QY 61 NCDKSLTLFGDKLCTVATIREYGENVADCCAKQEPERNECFLOHKKDDPNLPRLVRPEV 120  
 DB 83 NCDKSLTLFGDKLCAPLSLREHYGLADCCKEEPERNECFLOHKKNDPNIPKL-KPDE 141

QY 121 DVNCTAFHDNBTFLKKVLYEIAIRRHPPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 142 VALCADFOEDQKFWGKLYEIAIRRHPPYFAPPELLYYAIIYKDFSECCQADKAACLLP 201  
 QY 181 KDELDRDEKASSAKORLKCASLQKGFGERAFKAMAVARLSORFPAKAEFAEYSKLYDTLTK 240  
 DB 202 KIEHLREKVLISAARLKCASIQKGFGERAFKANSRLSQRFPKADFTISKIVTDLAK 261  
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKULKECEKPLLEKSHCIAEVENDEMPA 300  
 DB 262 VHKECHGDLLECADRADLAKYICENQDTISTKLKCECDKLEKSHCIAEAKDELP 321  
 QY 301 DIPSLAADFVSKDKVCKYAEAKDFLGMFLYEYARRHPDYSWLLLRLLAKIYETTLK 360  
 DB 322 DNLPHLEHDFVEKCKYKKEADKDFLGTFLYEYARRHPDYSVLLRIAKIYEATLEDC 381  
 QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLEGEYFQNALLVRYTKKVPQVST 420  
 DB 382 CAKEDPPACYATVDFKQPLVDPEKNLIKQNCLEFEKLEGEYFQNALIVRYTKKVPQVST 441  
 QY 421 PTLVEYSNLIKGVSKCKHPSAKMPCAEEDYLVVLNOLCVLHEKTPVSDRVTKCCTES 480  
 DB 442 PTLVEVARKLGLVSKCKRPEERUSCAEDYLSVLNOLCVLHEKTPVSEKVKCCTES 501  
 QY 481 LVNRRPCFSALVEYDITYPKFNAETFTFHADICTLSEKEROIKKQTAUVLVGHKPKAT 540  
 DB 502 LVNRRPCFSALTDETYKPEFVEGTFTHADICTLPEDEKQIKKQTAUVLVGHKPHAT 561  
 QY 541 KEQLKAVMDDFAAVFEKCKKADDCETCFABEKKLV 576  
 DB 562 EEQLTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 597  
 RESULT 11  
 ALBU\_MERUN STANDARD; PRT; 609 AA.  
 AC O35090;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB  
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 OC NCBI\_TaxID=10047;  
 [1]  
 RP SEQUENCE FROM N. A.  
 RC STRAIN=MGS IDR; TISSUE=Liver;  
 RX MEDLINE=98116663; PubMed=9455485;  
 RA Yoshida K., Seto-Ohashima A., Sinohara H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 RL DNA Res. 4:351-354(1997).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB006197; BAA21765.1; -;

[3] SEQUENCE OF 25-222.  
 RX MEDLINE=78109429; PubMed=564345;  
 RA Isemura S., Ikenaka T.;  
 RT "Amino acid sequences of fragments I and II obtained by cyanogen  
 bromide cleavage of rat serum albumin.";  
 RL J. Biochem. 83:35-48(1978).  
 [4] SEQUENCE OF 223-288 AND 572-608.  
 RX MEDLINE=76260153; PubMed=956149;  
 RA Isemura S., Ikenaka T.;  
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
 the amino acid sequences of four fragments.";  
 RL J. Biochem. 79:1183-1196(1976).  
 [5] SEQUENCE OF 166-174.  
 RX MEDLINE=79001617; PubMed=80265;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
 RL Cancer Res. 38:3483-3486(1978).  
 [6] FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC - FUNCTION: NRP regulates fat digestion, lipid absorption, and  
 blood flow (potential).  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - TISSUE SPECIFICITY: Plasma.  
 CC - SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC - SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; V01222; CAA24532.1; -  
 CC PIR; A93872; ABRTS.  
 CC HSP; P02768; 1E7B.  
 CC InterPro; IPR00264; Serum albumin.  
 CC Pfam; PF00273; transport\_prot; 3.  
 CC PRINTS; P00802; SERUMALBUMIN.  
 CC ProDom; PD002486; Serum\_albumin; 1.  
 CC SMART; SM00103; ALBUMIN; 3.  
 CC PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608 SERUM ALBUMIN.  
 FT PEPTIDE 166 174 NEUTROTENSIN-RELATED PEPTIDE.  
 FT DOMAIN 25 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.  
 FT DOMAIN 404 595 ALBUMIN 3.  
 FT METAL 27 27 COPPER.  
 FT DISULFID 77 86 BY SIMILARITY.  
 FT DISULFID 99 115 BY SIMILARITY.  
 FT DISULFID 114 125 BY SIMILARITY.  
 FT DISULFID 148 193 BY SIMILARITY.  
 FT DISULFID 192 201 BY SIMILARITY.  
 FT DISULFID 224 270 BY SIMILARITY.  
 FT DISULFID 269 277 BY SIMILARITY.  
 FT DISULFID 289 303 BY SIMILARITY.

PT DISULFID 302 313 BY SIMILARITY.  
 FT DISULFID 340 385 BY SIMILARITY.  
 FT DISULFID 384 393 BY SIMILARITY.  
 FT DISULFID 416 462 BY SIMILARITY.  
 FT DISULFID 461 472 BY SIMILARITY.  
 FT DISULFID 485 501 BY SIMILARITY.  
 FT DISULFID 500 511 BY SIMILARITY.  
 FT DISULFID 538 583 BY SIMILARITY.  
 FT DISULFID 582 591 BY SIMILARITY.  
 FT VARIANT 262 262 V -> L (IN REF. 5).  
 FT CONFLICT 174 174 Y -> L (IN REF. 5).  
 SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;  
 Query Match 78.2%; Score 2426; DB 1; Length 608;  
 Best Local Similarity 73.4%; Pred. No. 2.2e-148;  
 Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRPKDGLGHNFKALVLIAPQVLCQCPEDHVKLVNTEFAKTCAVDESAE 60  
 DB 25 EAHKSEIAHFKDGLGSHFGLVLIAPSQVLCQCPYEEHVKLVQEVTDFAKTCAVDESAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120  
 DB 85 NCDKSLHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLOHKDDNPNLPPORPEA 144  
 QY 121 DVMTAFHDNEETPLKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCOAAKACLLP 180  
 DB 145 EAMCTSFQENPTFSLGHYLEHVARHPYFYAPPELLFYAEKYNELVTCCTTESDKAACLT 204  
 QY 181 KLDELREGKASSAKQRLKASLQKQGERAFKAWAVARLSQRPFKAEFAEVSKLVDLT 240  
 DB 205 KLDVKEKALVAARVQRMKSSQWQGERAFKAWAVARMSQRPFAEFAEITKLATDVT 264  
 QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISKLKCECKPILKSHCHIAEVENDEMPA 300  
 DB 265 INKECCHGDLLECCADRAELAKYMCENQATISSKLOACCDKPVLOKSQCLAEHNDIPA 324  
 QY 301 DLPSLAADFVESDVCNVAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTVEITTEK 360  
 DB 325 DLPSIADFVEDKEVCNVAEAKDVFLGTFLYEVARRHPDYSVLLRLAKYEAITLEK 384  
 QY 361 CAADAPHECAKVEDEPKLVPEPNLIKONCELFQGLGKFNQALLVRYTKKVPQVST 420  
 DB 385 CAEGDPACVGTVLAEPQLVPEEPKLVKTNCELYEKLGEYGFQNAVLYRYTKAPQVST 444  
 QY 421 PTLVEVSRNLGKSGCKCKHPEAKRMPCAEYLSVNLQVLCVLEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEAARNLGRVGTCKCTLPEAQLPCVEDYLSAILNRLCVLHEKTPVSEKVTCCSGS 504  
 QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 540  
 DB 505 LVERRPCFSALTVDETVVPKEFKAETFTFHSIDICTLPDKEKQIKKQATALVELVHKPKAT 564  
 QY 541 KEQLKAVMDDFAAEVKECKKADDETCFAEKGKLVAAASQAL 583  
 DB 565 EDQLKTVMGDFAQFVDKCKCAADKDNCFATEGPNLVARSKEAL 607  
 RESULT 10  
 ID ALBU\_PIG STANDARD; PRT; 605 AA.  
 AC P08835; Q29018;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor (Fragment).  
 GN ALB.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```
RESULT 8
ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90098888; PubMed=2602160;
RA Brown W.M., Diegisielska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17055; CAA34903.1; -.
DR PIR; S06936; ABSH.
DR HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 147 192 BY SIMILARITY.
FT DISULFID 191 200 BY SIMILARITY.
FT DISULFID 223 269 BY SIMILARITY.
FT DISULFID 268 276 BY SIMILARITY.
FT DISULFID 288 302 BY SIMILARITY.
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 383 392 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;
```

```
Query Match 78.4%; Score 2432.5; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 8.4e-149;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;
QY 1 DAHSEVAVRPFKDLGEENFKALVLTAFQYLOCCPFEDHVKLVNTEFAKTCVADESAAE 60
DB 25 DTHKSEIAHRFNDLGEENFQGLVLTAFSQYLOCCPFEDHVKLVNTEFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPRLVRPEV 120
DB 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQPERNECFLNHKDDSPDKL-KPEP 143
QY 121 DVNCTAFHONETFLKKYLYETARRHPYFAPPELLFFAKRYKAAFTTECCQAADKACILP 180
DB 144 DTLCAEFKADKKFWGKLYEYARRHPYFAPPELLYYANKYNGVFOCCQAEDKGCILP 203
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240
DB 204 KIDAMREKVLASSARQRLCASIQKFGERALKXAVARLSQKFPKADFTDVKIVTDLT 263
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISGKLKECCKEKPLLEKSHGICAEVNDMPA 300
DB 264 VHKCCGDLLECCADRADLAKYICDHQDALSSKLKECCDKPYLEKSHGICAEVDKAVPE 323
QY 301 DLPLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLRAKTVETLEK 360
DB 324 NLPLTADFAEDKEVCNKYQEAQKDVGLSFLYESSRHPYAVSVLLRLAKEVEATLEDC 383
QY 361 CAAADPHECVAKVDFBPKPLVEEPQNLIKONCELPQOLGEYFQNALLYRYTKKVPQVST 420
DB 384 CAKEDPHACVATVDFDKLHLVDPEQNLIKONCELPQOLGEYFQNALLYRYTKKVPQVST 443
QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
DB 444 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 503
QY 481 LVNRRPCFSALYDEYTPVPRFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
DB 504 LVNRRPCFSDLTLDYTVVPKPFDEKFTFHADICTLPTDEKQIKKOTALVELVKKPKAT 563
QY 541 KEQIKAVMDFFAIVKECKKADDEKTCFAEKGKLVAAQAL 583
DB 564 DEQLKTYMENFVAFVDKCAADDEKGVLEGPRLVASTQAL 606
```

```
RESULT 9
ALBU RAT
ID ALBU_RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin
RT messenger RNA.";
RL J. Biol. Chem. 252:6846-6855(1977).
```

```

CC      ENEL; U18344; AAB56347.1; -.
DR      HSP; P02768; 1E7B.
DR      InterPro; IPR000264; Serum albumin.
DR      Pfam; PF00273; transport prot; 3.
DR      PRINTS; PR00802; SERUMALBUMIN.
DR      ProDom; PD002486; Serum albumin; 1.
DR      SMART; SM001103; ALBUMIN; 3.

```

[illegible]

RN [2] SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RP TISSUE=Liver;  
 RC Barry T., Power S., Gannon F.;  
 RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RC Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;  
 RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RL [4]  
 RN SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RP Wu H.T., Huang M.C.;  
 RA "The complete cDNA sequence of bovine serum albumin."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-32.  
 RX MEDLINE=80024278; PubMed=489109;  
 RA McGilivray R.T.A., Chung D.W., Davie E.W.;  
 RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin."  
 RL Eur. J. Biochem. 98:477-485(1979).  
 RN [6]  
 RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.  
 RA Brown J.R.;  
 RL "Structure of bovine serum albumin."  
 RT Fed. Proc. 34:591-591(1975).  
 RN [7]  
 RP REVISIONS TO 190-195.  
 RA Brown J.R.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [8]  
 RP SEQUENCE OF 402-433.  
 RX MEDLINE=82023364; PubMed=7283978;  
 RA Reed R.G., Putnam F.W., Peters T. Jr.;  
 RT "Sequence of residues 400-403 of bovine serum albumin."  
 RL Biochem. J. 191:867-868(1980).  
 RN [9]  
 RP SEQUENCE OF 19-28.  
 RX MEDLINE=77134075; PubMed=843354;  
 RA Patterson J.E., Geller D.M.;  
 RT "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin."  
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
 RN [10]  
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
 RX MEDLINE=91083649; PubMed=2260975;  
 RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;  
 RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS."  
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).  
 RN [11]  
 RP SEQUENCE OF 25-41.  
 RX MEDLINE=88267456; PubMed=3389500;  
 RA Heieh J.C., Lin F.P., Tam M.F.;  
 RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing."  
 RL Anal. Biochem. 170:1-8(1988).  
 RN [12]  
 RP SEQUENCE OF 437-451.  
 RA Vilbois F.;  
 RL Submitted (AUG-1998) to Swiss-Prot.  
 RN [13]  
 RP DISULFIDE BONDS.  
 RA Brown J.R.;  
 RL "Structure of serum albumin: disulfide bridges."  
 RT Fed. Proc. 33:1389-1389(1974).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- ALLERGEN: Causes an allergic reaction in human.  
 CC -1- SIMILARITY: Belongs to the ALB/AF/VPB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M73993; AAA51411.1; -;  
 DR EMBL; X58989; CAA41735.1; -;  
 DR EMBL; Y17769; CAA76847.1; -;  
 DR EMBL; AF542068; AAN17824.1; -;  
 DR HSSP; P02768; 1E7B;  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
 KW Polymorphism.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 607 SERUM ALBUMIN.  
 FT DOMAIN 25 204 ALBUMIN 1.  
 FT DOMAIN 211 396 ALBUMIN 2.  
 FT DOMAIN 403 594 ALBUMIN 3.  
 FT METAL 27 27 COPPER (BY SIMILARITY).  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 393 392  
 FT DISULFID 415 461  
 FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 FT VARIANT 214 214  
 FT CONFLICT 302 302 A -> T.  
 FT CONFLICT 304 305 C -> K (IN REF. 6).  
 FT CONFLICT 324 324 N -> D (IN REF. 6).  
 FT CONFLICT 394 395 ST -> TS (IN REF. 6).  
 FT CONFLICT 437 437 K -> R (IN REF. 12).  
 FT CONFLICT 493 494 SE -> ES (IN REF. 6).  
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;  
 Best Local Similarity 75.8%; Pred. No. 5.9e-150;  
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;  
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQYLCQCPFDHVKLVNEVTEFAKTCVADSSAE 60  
 Db 25 DTHKSEIHRFKDLGEEHFKGLVLIAPSVYLCQCPFDHVKLVNELTEFAKTCVADSSHA 84  
 QY 61 NCDKSLHTLFQDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDNDNPLPLVPEV 120  
 Db 85 GCEKSLHTLFQDELCKVASLRETYGDMADCCOEPERNECFLSHKDSDPLPKL-KPDP 143  
 QY 121 DVMCTATHDNETFLKKVLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 505 LVNRPFCPSGLVDEVTYVYKFNATFTFHADLCTLPFAEKQVKKQFALVELLKHKPKAT 564  
QY 541 KEQLKAVMDDPFAAFVEKCKADKCTCFABEGKLVAAASQAL 583  
Db 565 DBQLKTVMGDFGAFVEKCAENKEGCFSEBGPVLVAAASQAL 607

## RESULT 5

ALBU\_HORSE STANDARD; PRT; 607 AA.  
ID ALBU\_HORSE AC P35747;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Equ c 3).  
GN ALB.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
RC TISSUE=Liver;  
RX MEDLINE=93345495; PubMed=8344282;  
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
at 0.27-nm resolution."  
RL Eur. J. Biochem. 215:205-212 (1993).  
CC -I- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water. Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- TISSUE SPECIFICITY: Plasma.  
CC -I- ALLERGEN: Causes an allergic reaction in human. Binds IGE.  
CC -I- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -I- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X74045; CAA52194.1; -;  
CC PIR; S34053; ABHOS.  
CC HSSP; P02768; 1E7B.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; transport Prot; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC PRODOM; PD002486; Serum\_albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
KW SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607  
FT DOMAIN 25 204  
FT DOMAIN 211 396  
FT DOMAIN 403 594  
FT METAL 27 27  
FT METAL 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312

FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
SQ SEQUENCE 607 AA; 25656E830A1B90C5 CRC64;  
Query Match 79.8%; Score 2475.5; DB 1; Length 607;  
Best Local Similarity 76.3%; Pred. No. 1.5e-151;  
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;  
QY 1 DAHSEVAHRFDKLGHEENKALVLIAPAFQYLOCPEDHVKLVNVEYTEFAKTVADESAE 60  
Db 25 DTHKSEIAHRFDLGEKFKGLVLFVAFSQYLOCPEDHVKLVNVEYTEFAKCAADESAE 84  
QY 61 NCDKSLHTLFQDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
Db 85 NCDKSLHTLFQDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 143  
QY 121 DVMCTAFHNEETFLKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCCOAADAACALLP 180  
Db 144 DAQCAAFQEDPDKFLGKLYEVARRHYPFYGPPELLFHAEEYKADFTECCPADDKLACLIP 203  
QY 181 KLDELDRGKASSAKORLKASLQKFGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240  
Db 204 KLDALKERILLSSAKERLKCSSQNFGERAVKAVSWARLSQKPKADFAEVSKLVTDLTK 263  
QY 241 VHTCCGGDLLECCADDRADLAKYICENQDSISSKLECKECPLEKSHCIAEVENDEMPA 300  
Db 264 VHKCCGGDLLECCADDRADLAKYICENQDSISSKLECKECPLEKSHCIAEVENDEMPA 323  
QY 301 DLPSIADAFVESKDVCKNVAEKDVLGMFLYEYARRHPDYVSVLLRLAKTYETTLKXC 360  
Db 324 DLPALADAFADKKEICKHGKDAKDVFLGFLYEYARRHPDYVSVLLRLAKTYETTLKXC 383  
QY 361 CAAADPHECYAKVDFEKLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKAPQVST 420  
Db 384 CAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDFEEVGEYDFQNALIVRYTKKAPQVST 443  
QY 421 PTLVEVSNLGVSKCKCKHPEAKMPCAEEDYLSVLNQLCVLHEKTPVSDRVTKCTTES 480  
Db 444 PTLVEIGTLGVSRCCCKLPESRLPCSENLALNRLCVLHEKTPVSKITKCTDS 503  
QY 481 LVNRRPCPSALEVDETYVYKFNATFTFHADICTLSEKEQIKKQFALVELVXHKPKAT 540  
Db 504 LAERPCPSALELDEGVYVYKFNATFTFHADICTLSEKEQIKKQFALVELVXHKPKAT 563  
QY 541 KEQLKAVMDDPFAAFVEKCKADKCTCFABEGKLVAAASQAL 583  
Db 564 KEQLKAVMDDPFAAFVEKCKADKCTCFABEGKLVAAASQAL 606

## RESULT 6

ALBU\_BOVIN STANDARD; PRT; 607 AA.  
ID ALBU\_BOVIN AC P02769; C02787;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serum albumin precursor (Allergen Bos d 6).  
GN ALB.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;  
RL Submitted (AUG-1991) to the EMBL/GenBank/DBSJ databases.



DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serum albumin precursor (Allergen Can f 3).  
 GN ALB.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_TaxID=9615;  
 RN [1]\_TaxID=9615;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Beagle; TISSUE=Liver;  
 RA Hilger C.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=20148667; PubMed=10669848;  
 RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,  
 RA Valenta R., Spitzauer S.;  
 RT "Escherichia coli expression and purification of recombinant dog  
 RT albumin, a cross-reactive animal allergen.";  
 RL J. Allergy Clin. Immunol. 105:279-285(2000).  
 RN [3]  
 RP SEQUENCE OF 25-48.  
 RX MEDLINE=75011422; PubMed=4414013;  
 RA Dixon J.W., Sarkar B.;  
 RT "Isolation, amino acid sequence and copper(II)-binding properties of  
 RT peptide (1-24) of dog serum albumin.";  
 RL J. Biol. Chem. 249:5872-5877(1974).  
 RN [4]  
 RP SEQUENCE OF 25-38.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 RT dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 RN [5]  
 RP SEQUENCE OF 215-478 FROM N.A.  
 RC TISSUE=Salivary Gland;  
 RX MEDLINE=94201492; PubMed=7512102;  
 RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,  
 RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
 RT "Molecular characterization of dog albumin as a cross-reactive  
 RT allergen.";  
 RL J. Allergy Clin. Immunol. 93:614-627(1994).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human.  
 CC -!- SIMILARITY: Belongs to the ALB/AFB/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; AJ133489; CAB64867.1; -;  
 DR EMBL; Y17737; CAB76841.1; -;  
 DR EMBL; S72946; AAB30434.1; -;  
 DR HSP; P02768; IE7B.  
 DR HSC-2DPAGE; P49822;  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002466; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608  
 FT DOMAIN 25 205  
 FT DOMAIN 212 397  
 FT DOMAIN 404 595  
 FT METAL 27 27  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 148 193  
 FT DISULFID 192 201  
 FT DISULFID 224 270  
 FT DISULFID 269 277  
 FT DISULFID 289 303  
 FT DISULFID 302 313  
 FT DISULFID 340 385  
 FT DISULFID 384 393  
 FT DISULFID 416 462  
 FT DISULFID 461 472  
 FT DISULFID 485 501  
 FT DISULFID 500 511  
 FT DISULFID 538 583  
 FT DISULFID 582 591  
 FT CONFLICT 1 26  
 FT CONFLICT 146 146  
 FT CONFLICT 206 206  
 FT CONFLICT 349 349  
 FT CONFLICT 359 359  
 FT CONFLICT 448 448  
 FT CONFLICT 474 474  
 SQ SEQUENCE 608 AA; 86806 MW; 3CF1C8FF7DD8FC06 CRC64;  
 Query Match 82.6%; Score 2562; DB 1; Length 608;  
 Best Local Similarity 79.8%; Pred. No. 4.2e-157;  
 Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHFKDLGLENFALVLIAPAVQLQCPEDHVKLVNVEYTERAKTCAVESAE 60  
 DB 25 EAYKSEIAHRYNDLGEHFRGLVAFPSQYLOQCPEDHVKLAKVEYTERAKCAVESGA 84  
 QY 61 NCDKSLHTLPGDKLCTVATRETYGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEV 120  
 DB 85 NCDKSLHTLPGDKLCTVATRETYGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEV 144  
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFAKRYKAATFECQQAADKAACLLP 180  
 DB 145 DALCAAFQDNEQLFLGKYLVEIARRHPYFYAPELLFAKRYKAATFECQQAADKAACLLP 204  
 QY 181 KLDELREGKASSAKORLKCASLOKTEGAFKAWAVARLSQRPKAFKAVSVKLVDTLK 240  
 DB 205 KIEALREKVLSSAKERFKCASLOKTEGAFKAWAVARLSQRPKAFKAVSVKLVDTLK 264  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISKLKECCCKPLEKSHCHIAEVENEMPA 300  
 DB 265 VHKECCCHGDLLECCADRADLAKYICENQDSISKLKECCCKPLEKSHCHIAEVENEMPA 324  
 QY 301 DLPSLAADFVSKDVKCYKAEAKVFLGMLYEVARRHPDYVSVLLRLAKTYETLEK 360  
 DB 325 DLPSLAADFVSKDVKCYKAEAKVFLGMLYEVARRHPDYVSVLLRLAKTYETLEK 384  
 QY 361 CAAADPHCEKAVFDEPKLVVEEPQNLIKONCELFEOQGEYKTFQNALLVYTKKVPQVST 420  
 DB 385 CATDDPTCYAKVLDEFKPLVDFEQNLVKTNCBELFGLGEYGFQNALLVYTKKVPQVST 444  
 QY 421 PTLVEVRNLGKVGSKCKKPEAKRMPCADYLSVNLQVLHKEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEVRNLGKVGSKCKKPEAKRMPCADYLSVNLQVLHKEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

QY 301 DLPSLAADFYVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVWLLRLAKTYETTLK 360  
 DB 317 DLPSLAADFYVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVWLLRLAKTYETTLK 376  
 QY 361 CAADPHCYAKVDFEFKPLVEBPONLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420  
 DB 377 CAADPHCYAKVDFEFKPLVEBPONLIKONCELFQGLGEYKFNALLVRYTKVPQVST 436  
 QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLCLVHEKTPVSDRVTKCTES 480  
 DB 437 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLCLVHEKTPVSDRVTKCTES 496  
 QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
 DB 497 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQTALVELVKKPKAT 556  
 QY 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAL 583  
 DB 557 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAL 599

## RESULT 3

ALBU\_FELCA STANDARD; PRT; 608 AA.  
 ID ALBU\_FELCA  
 AC P49064; 1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serum albumin precursor (Allergen Fel d 2).  
 GN ALB.  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OC NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96194824; PubMed=8647469;  
 RA Hilger C., Grigioni F., Kohnen M., Hentges F.;  
 RT "sequence of the gene encoding cat (Felis domesticus) serum albumin.";  
 RL Gene 169:295-296(1996).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human.  
 CC -!- SIMILARITY: Belongs to the ALB/Arp/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announces/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X84842; CAA59279.1; -.  
 DR PIR; J04660; S57632.  
 DR HSRP; P02768; IEB7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 FT SIGNAL 1 18 By similarity.  
 FT PROPEP 19 24 By similarity.  
 FT CHAIN 25 608 SERUM ALBUMIN.  
 FT DOMAIN 25 205 ALBUMIN 1.

## RESULT 4

ALBU\_CANFA STANDARD; PRT; 608 AA.  
 ID ALBU\_CANFA  
 AC P49822; 077705; Q9TSZ4;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

FT DOMAIN 212 397  
 FT DOMAIN 404 595  
 FT METAL 27 27  
 FT BY SIMILARITY.  
 FT DISULFID 77 86  
 FT BY SIMILARITY.  
 FT DISULFID 99 115  
 FT BY SIMILARITY.  
 FT DISULFID 114 125  
 FT BY SIMILARITY.  
 FT DISULFID 148 193  
 FT BY SIMILARITY.  
 FT DISULFID 192 201  
 FT BY SIMILARITY.  
 FT DISULFID 224 270  
 FT BY SIMILARITY.  
 FT DISULFID 269 277  
 FT BY SIMILARITY.  
 FT DISULFID 289 303  
 FT BY SIMILARITY.  
 FT DISULFID 302 313  
 FT BY SIMILARITY.  
 FT DISULFID 340 385  
 FT BY SIMILARITY.  
 FT DISULFID 384 393  
 FT BY SIMILARITY.  
 FT DISULFID 416 462  
 FT BY SIMILARITY.  
 FT DISULFID 461 472  
 FT BY SIMILARITY.  
 FT DISULFID 485 501  
 FT BY SIMILARITY.  
 FT DISULFID 500 511  
 FT BY SIMILARITY.  
 FT DISULFID 538 583  
 FT BY SIMILARITY.  
 FT DISULFID 582 591  
 FT BY SIMILARITY.  
 SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;

Query Match 84.4%; Score 2620; DB 1; Length 608;  
 Best Local Similarity 82.0%; Pred. No. 8.1e-161;  
 Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRPKDILGEENFKALVLIATAQYLQCCFFEDHVKLVNNEVTEPAKTCVADESA 60  
 DB 25 EAHOSEIAHRRFNDEHFRGLVLFVAFQYLLQCCFFEDHVKLVNNEVTEPAKTCVADESA 84  
 QY 61 NCDXSLHTLFGDKLCTVATLEETYGEMADCCAKOEPERNECFLOHKDNDNPLVLVRPEV 120  
 DB 85 NCEKSLHELLGDKLCTVASLSDKYGEMADCCCKEPEERNECFLOHKDNDNPGQLVTPEA 144  
 QY 121 DVMCTAFHDNEETFLKVLVLIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 145 DAMCTAFHENEQRFLGKLYLVIARRHPYFYAPPELLFYAEYKGVFTECCAADKAACLLP 204  
 QY 181 KLDERDGGKASSAKORLKASLOKFGBERAFKAWAVARLSORFPKAFPAEYVKLVTDLT 240  
 DB 205 KVDALREKVLASSAKERLKCASLOKFGBERAFKANSVARLSQKFPKAFPAEYVKLVTDLAK 264  
 QY 241 VHTTECHGDLLECADDRADLAKYICENQDSISSKLKSCCEKPILEKSHCIAEVENDEMPA 300  
 DB 265 IHKECHGDLLECADDRADLAKYICENQDSISTKLKCCGKPVLEKSHCISEVERDELPFA 324  
 QY 301 DLPSLAADFYVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVWLLRLAKTYETTLK 360  
 DB 325 DLPLPLAVDFVEDKEVKQYQKADVFLGTFLYVEYSRRHPEYSVSLRLAKSEYATLEK 384  
 QY 361 CAADPHCYAKVDFEFKPLVEBPONLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420  
 DB 385 CATDDPPACYARHVFDEPKPLVEEPHNLVKTNCLEFELKELGEYGFQNALVRYTKVPQVST 444  
 QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLCLVHEKTPVSDRVTKCTES 480  
 DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLCLVHEKTPVSDRVTKCTES 504  
 QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
 DB 505 LVNRRPCFSALQVDETYVPKFNATFTFHADICTLPEAEKQIKKQKQSVALLKXKPKAT 564  
 QY 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAL 583  
 DB 565 BEQLKTVMGDFGFSVDKCCAAEDKEACFAEKGKLVAAASQAL 607



RC TISSUE=Liver, and Skeletal muscle;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Blakesley M., Madan A., Young A.C., Sherchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 25-609.  
 RX MEDLINE=76187907; PubMed=1225573;  
 RA Meloun B., Moravsek L., Kostka V.;  
 RT "Complete amino acid sequence of human serum albumin.";  
 RL FEBS Lett. 58:134-137(1975).  
 RN [9]  
 RP SEQUENCE OF 25-609.  
 RX Brown J.R., Shockley P., Behrens P.Q.;  
 RL (In) Bing D.H. (eds.);  
 RA The chemistry and physiology of the human plasma proteins, pp.23-40,  
 RL Pergamon Press, New York (1979).  
 RN [10]  
 RP SEQUENCE OF 1-455 FROM N.A.  
 RC TISSUE=Liver;  
 RA Menaya J., Farrilla R., Ayuso M.S.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [11]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=86140099; PubMed=24193129;  
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
 RT "The human albumin gene. Characterization of the 5' and 3' flanking  
 RT regions and the polymorphic gene transcripts.";  
 RL J. Biol. Chem. 261:3244-3251(1986).  
 RN [12]  
 RP SEQUENCE OF 222-229.  
 RX MEDLINE=76257808; PubMed=955075;  
 RA Walker J.E.;  
 RT "Lysine residue 199 of human serum albumin is modified by  
 RT acetylsalicylic acid.";  
 RL FEBS Lett. 66:173-175(1976).  
 RN [13]  
 RP SEQUENCE OF 25-44 AND 480-499.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [14]  
 RP DISULFIDE BONDS.  
 RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;  
 RT "Disulfide bonds in human serum albumin.";  
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
 RN [15]  
 RP BILIRUBIN-BINDING SITE.  
 RX MEDLINE=78186630; PubMed=656055;  
 RA Jacobsen C.;  
 RT "Lysine residue 240 of human serum albumin is involved in high-  
 RT affinity binding of bilirubin.";  
 RL Biochem. J. 171:453-459(1978).

RN [16]  
 RP VARIANT CANTERBURY ASN-337.  
 RX MEDLINE=87157744; PubMed=3828358;  
 RA Brennan S.O., Herbert P.;  
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
 RT domain of serum albumin.";  
 RL Biochim. Biophys. Acta 912:191-197(1987).  
 RN [17]  
 RP VARIANTS NAG-2 AND NAG-3.  
 RX MEDLINE=88068522; PubMed=3479777;  
 RA Takahashi N., Takanashi Y., Isobe T., Putnam F.W., Fujita M.,  
 RA Satoh C., Neel J.V.;  
 RT "Amino acid substitutions in inherited albumin variants from  
 RT Amerindian and Japanese populations.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
 RN [18]  
 RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
 RX MEDLINE=89345611; PubMed=2762316;  
 RA Arai K., Madison J., Huss K., Ishioke N., Satoh C., Fujita M.,  
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
 RT "Point substitutions in Japanese alloalbumins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
 RN [19]  
 RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
 RX MEDLINE=90115905; PubMed=2404284;  
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
 RT "Point substitutions in albumin genetic variants from Asia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
 RN [20]  
 RP DESCRIPTION OF VARIANT REDHILL.  
 RX MEDLINE=90115852; PubMed=2104980;  
 RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;  
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
 RT human serum albumin whose precursor has an aberrant signal peptidase  
 RT cleavage site.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
 RN [21]  
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
 RX MEDLINE=91062352; PubMed=2247440;  
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
 RA Watkins S., Putnam F.W.;  
 RT "Mutations in genetic variants of human serum albumin found in  
 RT Italy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
 RN [22]  
 RP VARIANT VENEZIA.  
 RX MEDLINE=91296740; PubMed=2068071;  
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
 RA Minchiotti L., Putnam F.W.;  
 RT "A donor splice mutation and a single-base deletion produce two  
 RT carboxyl-terminal variants of human serum albumin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
 RN [23]  
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
 RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
 RA Matsuda Y.-I., Anaki I., Putnam F.W.;  
 RT "Genetic variants of serum albumin in Americans and Japanese.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
 RN [24]  
 RP VARIANT CASEBROOK ASN-518.  
 RX MEDLINE=91316157; PubMed=1859851;  
 RA Peach R.J., Brennan S.O.;  
 RT "Structural characterization of a glycoprotein variant of human serum  
 RT albumin: albumin Casebrook (494 Asp-->Asn).";  
 RL Biochim. Biophys. Acta 1097:49-54(1991).  
 RN [25]  
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
 RX MEDLINE=92190239; PubMed=1347703;  
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
 RA Rochu D., Porta F.;  
 RT "Two alloalbumins with identical electrophoretic mobility are produced

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 42.6731 Seconds  
(without alignment)  
713.823 Million cell updates/sec

Title: US-09-832-929-18  
Perfect score: 3103  
Sequence: 1 DAHSEVAHFVDLGEENFK.....TCFAEEGKLVASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 3103   | 100.0       | 609    | 1     | ALBU HUMAN  |
| 2          | 2942   | 94.8        | 600    | 1     | ALBU MACMU  |
| 3          | 2620   | 84.4        | 608    | 1     | ALBU FELCA  |
| 4          | 2562   | 82.6        | 608    | 1     | ALBU CANFA  |
| 5          | 2475.5 | 79.8        | 607    | 1     | ALBU HORSE  |
| 6          | 2450.5 | 79.0        | 607    | 1     | ALBU BOVIN  |
| 7          | 2446   | 78.8        | 608    | 1     | ALBU RABIT  |
| 8          | 2432.5 | 78.4        | 607    | 1     | ALBU SHEEP  |
| 9          | 2426   | 78.2        | 608    | 1     | ALBU RAT    |
| 10         | 2411.5 | 77.7        | 605    | 1     | ALBU PIG    |
| 11         | 2387   | 76.9        | 609    | 1     | ALBU MERUN  |
| 12         | 2378   | 76.6        | 608    | 1     | ALBU MOUSE  |
| 13         | 1557.5 | 50.2        | 615    | 1     | ALBU CHICK  |
| 14         | 1253.5 | 40.4        | 609    | 1     | FETA PANTR  |
| 15         | 1249.5 | 40.3        | 609    | 1     | FETA HUMAN  |
| 16         | 1242.5 | 40.0        | 609    | 1     | FETA GORGO  |
| 17         | 1203   | 38.8        | 607    | 1     | ALB2 XENLA  |
| 18         | 1200   | 38.7        | 609    | 1     | FETA HORSE  |
| 19         | 1164.5 | 37.5        | 606    | 1     | ALB1 XENLA  |
| 20         | 1084   | 34.9        | 605    | 1     | FETA MOUSE  |
| 21         | 1067   | 34.4        | 611    | 1     | FETA RAT    |
| 22         | 1055   | 34.0        | 599    | 1     | AFAM HUMAN  |
| 23         | 944    | 30.4        | 611    | 1     | AFAM MOUSE  |
| 24         | 928    | 29.9        | 608    | 1     | AFAM RAT    |
| 25         | 747.5  | 24.1        | 608    | 1     | ALB2 SALSA  |
| 26         | 742.5  | 23.9        | 608    | 1     | ALB2 SALSA  |
| 27         | 699    | 22.5        | 382    | 1     | ALBU RANCA  |
| 28         | 440.5  | 14.2        | 1423   | 1     | ALBU PETMA  |
| 29         | 386    | 12.4        | 476    | 1     | VTDB HUMAN  |
| 30         | 381    | 12.3        | 474    | 1     | VTDB RAT    |
| 31         | 378    | 12.2        | 476    | 1     | VTDB RABIT  |
| 32         | 372    | 12.0        | 476    | 1     | VTDB MOUSE  |
| 33         | 151.5  | 4.9         | 1605   | 1     | RRB1 MOUSE  |

## ALIGNMENTS

## RESULT 1

| ID | ALBU_HUMAN   | STANDARD | PERT | 609 AA |
|----|--|----------|------|--------|
| AC | P02768; O95574; Q13140; Q9P157; Q9UHS3; Q9UJZ0;                        |          |      |        |
| DT | 21-JUN-1986 (Rel. 01, Created)   |          |      |        |
| DT | 01-APR-1990 (Rel. 14, Last sequence update)                            |          |      |        |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update)                          |          |      |        |
| DE | Serum albumin precursor.   |          |      |        |
| GN | ALB.   |          |      |        |
| OS | Homo sapiens (Human).  |          |      |        |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |          |      |        |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.             |          |      |        |
| OX | NCBI_TaxID=9606;   |          |      |        |
| RN | [1]  |          |      |        |
| RP | SEQUENCE FROM N.A.   |          |      |        |
| RX | MEDLINE=86196112; PubMed=3009475;                                      |          |      |        |
| RA | Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W., |          |      |        |
| RA | Beattie W.G., Dugaiczky A.;  |          |      |        |
| RT | "Molecular structure of the human albumin gene is revealed by          |          |      |        |
| RT | nucleotide sequence within q11-22 of chromosome 4.";                   |          |      |        |
| RL | J. Biol. Chem. 261:6747-6757(1986).                                    |          |      |        |
| RN | [2]  |          |      |        |
| RP | SEQUENCE FROM N.A., AND VARIANT LYS-420.                               |          |      |        |
| RX | MEDLINE=82081882; PubMed=6171778;                                      |          |      |        |
| RA | Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,             |          |      |        |
| RA | Najarian R.C., Seeburg P.H., Wion K.L.;                                |          |      |        |
| RT | "The sequence of human serum albumin cDNA and its expression in E.     |          |      |        |
| RT | coli.";  |          |      |        |
| RL | Nucleic Acids Res. 9:6103-6114(1981).                                  |          |      |        |
| RN | [3]  |          |      |        |
| RP | SEQUENCE FROM N.A., AND VARIANT GLY-121.                               |          |      |        |
| RX | MEDLINE=82105994; PubMed=6275391;                                      |          |      |        |
| RA | Dugaiczky A., Law S.W., Dennison O.E.;                                 |          |      |        |
| RT | "Nucleotide sequence and the encoded amino acids of human serum        |          |      |        |
| RT | albumin mRNA.";  |          |      |        |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).                          |          |      |        |
| RN | [4]  |          |      |        |
| RP | SEQUENCE FROM N.A.   |          |      |        |
| RC | TISSUE=Liver;  |          |      |        |
| RA | Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;                               |          |      |        |
| RL | Submitted (SRP-1999) to the ENBL/GenBank/DBJ databases.                |          |      |        |
| RN | [5]  |          |      |        |
| RP | SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).          |          |      |        |
| RC | TISSUE=Fetal liver;  |          |      |        |
| RA | Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,    |          |      |        |
| RA | Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;             |          |      |        |
| RT | "Functional prediction of the coding sequences of 121 new genes        |          |      |        |
| RT | deduced by analysis of cDNA clones from human fetal liver.";           |          |      |        |
| RL | Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.                |          |      |        |
| RN | [6]  |          |      |        |
| RP | SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.                   |          |      |        |
| RA | Huang M.C., Wu H.T.;   |          |      |        |
| RT | "The cDNA sequences of human serum albumin.";                          |          |      |        |
| RL | Submitted (AUG-2002) to the ENBL/GenBank/DBJ databases.                |          |      |        |
| RN | [7]  |          |      |        |
| RP | SEQUENCE FROM N.A.   |          |      |        |

O8nf91 homo sapien  
Q9p2e9 homo sapien  
Q08696 drosophila  
Q13439 homo sapien  
P35749 homo sapien  
P49454 homo sapien  
Q58718 methanococc  
P35748 oryctolagus  
Q90988 gallus gall  
Q14789 homo sapien  
O75962 homo sapien  
P25386 saccharomyc

34 144.5 4.7 8797 1 SNE1\_HUMAN  
35 138.5 4.5 1410 1 RRB1\_HUMAN  
36 133.5 4.3 1391 1 MST2\_DROHY  
37 132.5 4.3 2230 1 GOA4\_HUMAN  
38 129.5 4.2 1972 1 MYH8\_HUMAN  
39 129 4.2 3210 1 CENF\_HUMAN  
40 128 4.1 1005 1 RA50\_METJA  
41 126.5 4.1 1972 1 MYH8\_RABIT  
42 126 4.1 1189 1 SMCT2\_CHICK  
43 126 4.1 3259 1 GOB1\_HUMAN  
44 125 4.0 3038 1 TRIO\_HUMAN  
45 124.5 4.0 1790 1 USO1\_YEAST

Db 141 VEPVTSCEAYBEDRETWNKFIYBIARRHPLXYAPTILLWAARYDKIIPSCCKAENAVE 200  
Qy 177 CLLPKLDLDEKSGASSAKRLKASLOKFGERAFAKAVARLSORPKAFSAFVSKLVT 236  
Db 201 CFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFOAITVTKLSOKFTKVNFTSIQKLV 260  
Qy 237 DLTQVHTECHCHDLLECCADRDALAKYICENQDSISSKLKCCSKPLLEKSHCIAEVEND 296  
Db 261 DVAVHVEHCRCRQDVLCDLQDGEKIMSYICSQDQTLNSNITECKTLTLERQOCIHAEND 320  
Qy 297 ENPADLPSLAADPVESKDVCKQYAAKQVFLGMFLYEVARRHPDYSVLLRLAKTYETT 356  
Db 321 EXPEGLSPNLNRLFDGNDPNSGGEKNIIFLASFVHEYSRRHPQLAVSVILRVKAGQOEL 380  
Qy 357 LEKCCAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOIGYKFNALLVRYTKVP 416  
Db 381 LEKCCQTEPNLSCQDGEELQYIOESQALAKRSQGLFQKLGYYLQNAFLVATKKAP 440  
Qy 417 QVSTPTLVEVSNLKGKCKHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKC 476  
Db 441 QLTSSLMATRKMAATAATCCQLEDKLLACGEAADIIGHLCIRHEMTVPNPGVGQC 500  
Qy 477 CTESLVNRRPCSALEVDVTPYKFNABTFTFHADICTLSEKQIKKQATALVELVGHK 536  
Db 501 CHSSVARRPCSSLVVDVTPYPPAFSDDKFIFHKDLCOAQVQALQTKWQEFNLNVKQX 560  
Qy 537 PRATKEQLKAVMDDFAAFVEKCKCKADDETCFAEKGKLVAAASQAALGL 585  
Db 561 PQITEQLETVIADFSGILEKCCQGEQVCFABEGCKLISKTRTALGV 609

RESULT 15  
ABXL72  
74K albumin precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: B41682; S02693; A05288  
R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.  
Mol. Endocrinol. 3, 464-473, 1989  
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.  
A:Reference number: A41682; MUID:89313788; PMID:2747653  
A:Accession: B41682  
A:Molecule type: mRNA  
A:Residues: 3-607 <SCH>  
A:Cross-references: GB:M21442; NID:G213930; PIDN:AAA49637.1; PID:G213931  
R:Schorpp, M.; Doebeiling, U.; Wagner, U.; Ryffel, G.U.  
J. Mol. Biol. 199, 83-93, 1988  
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deleted exon 1.  
A:Reference number: S02692; MUID:88172470; PMID:2451026  
A:Accession: S02693  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-48 <SCH>  
A:Cross-references: EMBL:Z26826  
R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J. Eur. J. Biochem. 146, 489-496, 1985  
A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization of the 74-kDa albumin mRNA.  
A:Reference number: A05288; MUID:85126974; PMID:3971963  
A:Accession: A05288  
A:Molecule type: mRNA  
A:Residues: 459-502, 'L', 504-557 <WOL>  
A:Cross-references: GB:M28276  
A:Note: the authors translated the codon TAT for residue 63 as Thr  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyrones.  
C:Genetics: 27/1  
A:Introns: 27/1  
C:Superfamily: serum albumin; serum albumin repeat homolog  
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>

F:25-607/Product: 74K serum albumin #status predicted <MAT>  
F:32-201/Domain: serum albumin repeat homolog <SA1>  
F:220-393/Domain: serum albumin repeat homolog <SA2>  
F:412-591/Domain: serum albumin repeat homolog <SA3>  
F:30/Binding site: copper (His) #status predicted  
F:80-88/101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392  
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.8%; Score 1205; DB 1; Length 607;  
Best Local Similarity 39.3%; Pred. No. 3.3e-72;  
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

Qy 3 HKSEVAHRFKDLGEENFKALVLIAPQYLOCCPFEDHVKLVNEVTEFAKTCVADGEAENC 62  
Db 30 HKKHADVYALTERTFKGTJLAIVSQNLQKSELSKLVNEINDFAKSINDKTIPE-C 88  
Qy 63 DKSLHILFGKLCITVATLRETYGEMADCCAKQBPENECFLQHKDNPPLPRVREVDV 122  
Db 89 EKPVGILFFDKLCADPAVGVNYSKECCAKQDPERAQCFKAHRDHEHT--SIKPEPEE 145  
Qy 123 MCTAFHDNETFLKVLVEIARRHPYFAPPELLFPAKRYKAAFTCCQADKAACLLPKL 182  
Db 146 TCCLKKEHPDDLISAFIHEARNHPDLYPPAVLALTQVHKLAHCCCEDEKCKCFSEKM 205  
Qy 183 DELRDECKASSAKQRLKCAQLQKFGERAFAKAVARLSORPKAFSAFVSKLVTDLTKVH 242  
Db 206 KQLMKQSHSIEDQHHFCWILDNFPKVLKALNARVSHYPAKFLAHNFTVEVTHFI 265  
Qy 243 TECCHGDLLECADRDALAKYICENQDSISSKLKCCSKPLLEKSHCIAEVENDMPADL 302  
Db 266 KDCCHDDMFECMTERTLEHTCQHKDELSSKLEKCCNIPLEERTYCIVTLENDVPAEL 325  
Qy 303 PSLAADPVESKDVCKQYAAKQVFLGMFLYEVARRHPDYSVLLRLAKTYETTLEKCCA 362  
Db 326 SQPITEFTEDPHVCKEYAEENVEFLGYLHVSERKQELSEQFLQSAKESYLKCKCK 385  
Qy 363 AADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOIGYKFNALLVRYTKVPQVSTPT 422  
Db 386 TDNPECYKQDADRFMNEAKERFAYLKQNCIDLHEHGEYLFENELLRYTKQMPQVSDET 445  
Qy 423 LVEVSNLKGKCKCKHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCTESLV 482  
Db 446 LIGAHQVADIGEHCACAVPENQRMPCAEGLDILIGKQERQKTFINNVAHCCIDSYS 505  
Qy 483 NRRPCFSALEVDVTPYKFNABTFTFHADICTLSEKQIKKQATALVELVGHKPKATKE 542  
Db 506 GMRSCFTALGDEDDYVPPVTDDTFHDDKICTANDKEKHQIKQKFLVLIKIVSPKLEKN 565  
Qy 543 QLKAVMDDFAAFVEKCKCKADDETCFAEKGKLVAAASQ 580  
Db 566 HIDECSAEFLKQVQKCCCTADEHQPCFDTKEPVLIEHCQ 603

Search completed: April 19, 2004, 12:02:19  
Job time : 73.3019 secs

A;Residues: 19-45;63-97;102-107;122-184;187-249;255-489;507-609 <REC>  
R;Yachnin, S.; Heu, R.; Heinrikson, R.L.; Miller, J.B.  
Biochim. Biophys. Acta 493, 418-428, 1977  
A;Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric  
A;Reference number: A90624; MUID:77242506; PMID:70228  
A;Accession: A90624  
A;Molecule type: protein  
A;Residues: 'S',20-22,'S',24-35 <YAC>  
A;Note: dimeric and trimeric forms have been found in addition to the monomeric form  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 37, 3663-3667, 1977  
A;Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and  
A;Reference number: A90757; MUID:78001760; PMID:71198  
A;Accession: A90757  
A;Molecule type: protein  
A;Residues: 'S',20-30,'A',32-37,'A' <AOY>  
R;Ruoholahti, E.; Pihko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Konttinen, A.  
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974  
A;Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains  
A;Reference number: A93042; MUID:75018719; PMID:4138095  
A;Accession: A93042  
A;Molecule type: protein  
A;Residues: 'S',20-24,'Q',26-30,'A',32-35,'E',37-39 <RUO>  
R;Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamaoki, T.  
J. Biol. Chem. 260, 5055-5060, 1985  
A;Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking region  
A;Reference number: A92520; MUID:85182629; PMID:2580830  
A;Contents: annotation; Gene, exons and introns  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A;Reference number: A90758; MUID:79001617; PMID:80265  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 39, 3571-3574, 1979  
A;Title: Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability  
A;Reference number: A90759; MUID:80001710; PMID:89900  
A;Contents: annotation; bilirubin binding  
C;Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma of  
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AFP  
C;Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin  
Properties.  
C;Genetics:  
A;Gene: GDB:AFP  
A;Cross-references: GDB:119660; OMIM:104150  
A;Map position: 4q11-4q13  
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-609/Product: alpha-fetoprotein #status experimental <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;22/Binding site: copper (His) #status experimental  
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472  
F;249/Binding site: bilirubin (Lys) #status predicted  
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 40.3%; Score 1249.5; DB 1; Length 609;  
Best Local Similarity 39.9%; Pred. No. 3.7e-75;  
Matches 235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;  
QY 3 HKSE-----VAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADE 57  
DB 22 HRNEYGIASILDYSOCTAIEISLADLTIFFAQFVQEAITYKEVSKVMDALTAIEKPTGDE 81  
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHQDNPN-NLPLRV 116  
DB 82 QSSGLENQLPAFLEELCHEKEILEKYGH-SDCCSQSSEGRHNCFLAHHKPTPASIFLQ 140  
QY 117 RPEVDVMTAFHDNETFLKYLVIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAA 176

Db 141 VBPVTSCEAYEDRETWNKFIYEIARRHPFLYAFATILLWAARYDKIIPSCCAENAVE 200  
QY 177 CLIFPKDLDEKASSAKORLKCASLQKFGEPAPKAWAVARUSQFFPKAFAEVSKLYT 236  
Db 201 CFQTKAATVTKELRESSLLNQHACAVMKNFGRFTFOAITVTKLSQKFTKVNFTIQLVL 260  
QY 237 DLTQVHTCECHGDLLECCADRADLAKYICENQDISSKLKECCPKLLEKSHCIAEVEND 296  
Db 261 DVAHVHCHRCRGVDDCLQGEKIMSYICQQDTLSNKTCCCKLTTLERGQCIHAEND 320  
QY 297 EMPADLPSLAADPVESKDVCKYAEAKDVLGMLFLEYARRHPDYSVLLRLAKYETT 356  
Db 321 EKPEGLSPNLNRFLGDRDFNQFSSEKNIPLASVFVHEYSRRHPQLAVSVILRVAKGQEL 380  
QY 357 LEKCCAAADPHECVAKVDFEFKPLVEPQNLIKQNCLEFQELGEYKFTONALLVYTKKVP 416  
Db 381 LEKCFQTEPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGVEYVYQNAFLVAYTKAP 440  
QY 417 QVSTPTLVEVSRNLGVKSGKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKC 476  
Db 441 QLTSSSELMAITRKMAATAATCCQLSEDKLLACGSGAADIIGHLCIRHEMTVPNPGVGQC 500  
QY 477 CTESLVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKERQIKKQATALVELVKHK 536  
Db 501 CTSSYANRRPCFSLVVDVTVPPAFSDDKFIHKDJCOAGGVALQTMKQEFLLNLVKQK 560  
QY 537 PKATKEQLKAVMDFAAFVKCKCCKADDKTCFABEGKLVAAQAALGL 585  
Db 561 PQITEQLEAVIADFSGLEKCCQEQEVCFABEGQKLSKTRAAALGV 609  
RESULT 14  
PFGO  
alpha-fetoprotein precursor - gorilla  
C;Species: Gorilla gorilla (gorilla)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C;Accession: A37970  
R;Ryan, S.C.; Zielinski, R.; Dugaiczky, A.  
Genomics 9, 60-72, 1991  
A;Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primate  
A;Reference number: A37970; MUID:91169517; PMID:1706310  
A;Accession: A37970  
A;Molecule type: DNA  
A;Residues: 1-609 <RYA>  
A;Cross-references: GB:M38272; NID:9817963; PIDN:AAA73520.1; PID:gl77041  
C;Genetics:  
A;Map position: 4q11-12  
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 51  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-609/Product: alpha-fetoprotein #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;22/Binding site: copper (His) #status predicted  
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-4  
F;249/Binding site: bilirubin (Lys) #status predicted  
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 40.0%; Score 1242.5; DB 1; Length 609;  
Best Local Similarity 39.6%; Pred. No. 1.1e-74;  
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;  
QY 3 HKSE-----VAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADE 57  
DB 22 HRNEYGIASILDYSOCTAIEISLADLTIFFAQFVQEAITYKEVSKVMDALTAIEKPTGDE 81  
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHQDNPN-NLPLRV 116  
Db 82 QSSGLENQLPAFLEELCHEKEILEKYGH-LSDCSQSSEGRHNCFLAHHKPTPASIFLQ 140  
QY 117 RPEVDVMTAFHDNETFLKYLVIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAA 176

```

Db      330 LPSLVEKYIEDVEKCSFAGHDAMFASVFEVYSRRHDFEFTQILIMRIAKGYESLLEKCC 389
QY      362 AAADPHECYAKVDFEFLPVEEPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVSTP 421
Db      390 KTDNPAECYANAQEQNLQHIKETQDVVTKNCDLLHDGAEADFLKSILIRYTKQMPQVPTD 449
QY      422 TLIVESRNLGVKSGCKHPKRAKMPCAEDYLSVLNLQVLHETKTPVSDVTKCTESL 481
Db      450 LLEETGKMTTIGTKCCQJGDRMACSEGYLSIVIHDTCKQETTPINDNVQCCSOLY 509
QY      482 VNRRFCFSALEYDETVVPKFEAETTFHADICTLSEKERQIKKOTALVELVKKPKATK 541
Db      510 ANRPPCTAMGVDTKVPFPFNDPMSFDEKLCSPASEREGVMKLLINLIKRPQMT 569
QY      542 EQLKAVMDPFAFVEKCKKADDKETCFABEGKLVAAQAAAGL 585
Db      570 EQIKTIADGFTAMVCKKQSDINTCFGEEGANLVQSRATLGI 613

```

## RESULT 12

```

JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GB:U21916; NID:G841311; PIDN:AAA91641.1; PID:G841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
o similar properties and structure.
C:Genetics:
A:Gene: atp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <NAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.1%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;
QY 3 HKSE-----VAHFRKDLGEENFKALVLIIPAQYLOQCPPEFHVKLVNVEYTFKACTVADE 57
Db      22 HRNEYGLASILDVQCTAEINTDLATIFFAQVQBATYKSVKVKDALTAIEKPTGDE 81
QY 58 SAENCYKSLHTLFGDKLCTVATIREYTGEMADCAQEBERNECFLOHDDNP-NLPLRV 116
Db      82 QSAGCUENQPAFLLEELCREKEILEKYGH-SDCCSQSEGRHNCFLAHHKPTFASIPFFQ 140
QY 117 RPEVDVMTAFHNEETFLKLYIEARRHPFYAPPELLFAKRYKAAFTCCQAAADKAA 176
Db      141 VPEPVTSCAYEEDREFTNKFIYIARRHPFLYAPTILLWAARYDKIIPSCCKAENAVE 200
QY 177 CLIPKLDLDEDEKASSAKORLKASLQKGEGRAPKAVARLSQRPFAEFAEVSKLVT 236
Db      201 CFQTKATVTKELRESLLNHQACAVMKNFGTTFQAITVTKLSQKFTKVFTEIQKVL 260
QY 237 DLTVHTECGHDLLECADDRADLAKYICENQDISSKLKECCCKPDLLEKSHCIAEVND 296
Db      261 DVAVHVEHCRRGVLDCLQGEKIMSYICQQDTLSNKITECCKLTILRGQCIIHAEND 320
QY 297 EMPADLPSLAADFVESKDVCKVNAEKDVFGLMFLYEARRRHPDYSVLLLLRAKYETT 356

```

```

Db      321 EKTEGSLPNLNRFLGDRDNQPSGSEKNIFLASFVHEYSRHPQLAVSVILRAKGYQEL 380
QY      357 LEKCCAAADPHECYAKVDFEFLPVEEPQNLIKONCELFQQLGEYKFNALLVRYTKKVP 416
Db      381 LEKCFOTENLEQCDQGEELQYIQESQALAKRSCGLFQKLGEYFLQNAFLVAYTKAP 440
QY      417 QVSTPTLVEVSRNLGVKSGCKHPKRAKMPCAEDYLSVLNLQVLHETKTPVSDVTKTC 476
Db      441 QLTSSELMATRMAATAATCCQLSDKLLACGEGADIIIGHLCIRHETTFPNPVGQC 500
QY      477 CTSLVNRRFCFSALEYDETVVPKFEAETTFHADICTLSEKERQIKKOTALVELVKK 536
Db      501 CTSSYANRRFCFSSLVVDETVVPFASDDKFIHKDLCOAQGVALQTMKQEFNLINLVKQ 560
QY      537 PKATKEQLKAVMDPFAFVEKCKKADDKETCFABEGKLVAAQAAAGL 585
Db      561 PQITEGLEAVIADFGSLLEKCCQGEQVCFABEGQKLISKTRALLGV 609

```

## RESULT 13

```

PPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1339-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complet
A:Reference number: A26624; MUID:87195438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GIB>
A:Cross-references: GB:M6110; NID:G773678; PIDN:AAB58754.1; PID:G178236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krumlat
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein ge
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:Z19532; NID:G28527; PIDN:CAA79592.1; PID:G28528
A:Note: the authors translated the codon TAT for residue 26 as Thr
R:Moringa, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GB:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R:Beattie, W.G.; Dugaiczky, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequen
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 423-556 <BEA>
A:Cross-references: GB:J00076
R:Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Terrar
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45;160-97;102-107;122-184;187-249;255-489;507-609 <PUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly reprc
A:Reference number: A61480; MUID:91225826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```



```
QY 422 TLVEVSRNLGKVGSKCKRHPKARMPCAEDYLSVVLNQLCVLHKPTVPSDRVTKCTESL 481
DB 447 TLVEAARSILGRVTHCCALPEKRLPCVEDYLSAILNRVCLLHKTVPSEQVTKCCSGSL 506
QY 482 VNRPCFSALEVDVETVPKFAEFTFHADICTLSEKROIKKOTALVELVKKPKATK 541
DB 507 VERRPCFSALPVDVETVPKFAEFTFHADICTLSEKROIKKOTALVELVKKPKATK 566
QY 542 EQLKAVMDPFAAFVEKCCADKDKETCFABEGKKLVAASQAAL 583
DB 567 EQLKAVMDPFAAFVEKCCADKDKETCFABEGKKLVAASQAAL 608
RESULT 10
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; 148638
R:Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A>Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A:Reference number: A93055; MUID:88216123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <MIN>
A:Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Boccacchio, C.; Deschatrete, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A>Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A:Reference number: 148638; MUID:90269606; PMID:1971802
A:Accession: 148638
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:335-453/Domain: serum albumin repeat homology (fragment) <SA3>
Query Match 50.0%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 9.4e-116;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;
QY 75 CTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRVDMCTAFHNEETP 134
DB 1 CAIENLRENYGELADCTCKQEPERNECFLOHKDDNPNLPPFERPEAEAMCTSKENPTTF 60
QY 135 LKCYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLPKLDELDEGKASSA 194
DB 61 MGHYLEVARRRHPYFYAPPELLFYAEQYNEILTOCCABADKESCLTPKLDGVKEKALVSV 120
QY 195 KORLKASLOKFGERAFAKAVARLSORFPKAEVSKLAVTLTKVHTCCCHGDLLECA 254
DB 121 RQRVKCSMQKFGERAFAKAVARLSOTFFNADPAETIKLATDLTKVKKCCCHGDLLECA 180
QY 255 DDRADLAKYICENDDSISKLKCECEKPLLEKSHCIAEVENDEMPADLPISLAADPVESKD 314
DB 181 DDRAELAKYMCENQATISSKLQTCDDKPLKKAHCLSEVEHDTMPADLPISLAADPVESKD 240
QY 315 VCKRYAEAKQVFLGFMFLYEARRHPDYSVLLLLAKTYETLEKCCAAADPHECYAKVF 374
DB 241 VCKRYAEAKQVFLGFMFLYEARRHPDYSVLLLLAKTYETLEKCCAAADPHECYAKVF 300
QY 375 DEFKPLVEEPONLAKQNCLEPQOIGYKFNQNALVRYTKVPQVSTPTLVEVSRNLGKVG 434
DB 301 AEFQPLVEEPONLAKQNCLEPQOIGYKFNQNALVRYTKVPQVSTPTLVEVSRNLGKVG 360
QY 435 SKCCGHPKARMPCAEDYLSVVLNQLCVLHKPTVPSDRVTKCTESLWNRPCFSALEVD 494
DB 361 TKCCTLPDQRLPCVEDYLSAILNRVCLLHKTVPSEQVTKCCSGSLVERRPCFSALTVD 420
```

```
QY 495 ETVVPKFEFNAETFTFHADICTLSEKROIKKOT 527
DB 421 ETVVPKFEFNAETFTFHADICTLSEKROIKKOT 453
```

## RESULT 11

ABCHS

serum albumin precursor - chicken

C:Species: Gallus gallus (chicken)

C&gt;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999

C:Accession: S15571; A05078; A13451

R:Casady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.

submitted to the EMBL Data Library, July 1991

A:Reference number: S15571

A:Accession: S15571

A:Molecule type: mRNA

A:Residues: 1-615 &lt;CAS&gt;

A:Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748

R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, P.C.K.; Deeley, R.G.

J. Biol. Chem. 258, 4556-4564, 1983

A&gt;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein

A:Reference number: A05078; MUID:83161037; PMID:6187737

A:Accession: A05078

A:Molecule type: DNA

A:Residues: 1-28 &lt;HAC&gt;

A:Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039

R:Rosen, A.M.; Geller, D.M.

Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977

A&gt;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.

A:Reference number: A13451; MUID:78019943; PMID:911137

A:Accession: A13451

A:Molecule type: protein

A:Residues: 19-23 'M', 25-30 &lt;ROS&gt;

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,

mones (weak bonds with these hormones promote their transfer across the membranes), the

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-26/Domain: propeptide #status predicted &lt;PRO&gt;

F:27-613/Product: serum albumin #status predicted &lt;MAT&gt;

F:32-206/Domain: serum albumin repeat homology &lt;SA1&gt;

F:225-398/Domain: serum albumin repeat homology &lt;SA2&gt;

F:417-596/Domain: serum albumin repeat homology &lt;SA3&gt;

F:30/Binding site: copper (His) #status predicted

F:80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397

Query Match 50.2%; Score 1557.5; DB 1; Length 615;

Best Local Similarity 46.7%; Pred. No. 1.6e-95;

Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;

QY 3 HKSEVNRHFKDGLGENFKALVLIAPQYLQCPEDHVKLVNVEYTFKACTVADESANVC 62

DB 30 HKSEIARHNDLKEETFRKAVAMITPAQYLQRCSEGLSKVYKDVVDLQACKVANEDAPEC 89

QY 63 DKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPR-LVRPEVD 121

DB 90 SKPLPSIILDEICQVEKLRDYSYGAVADCCSKADPERNECFLSFKVQSQDFVQYQRPASD 149

QY 122 VNCTAFPHNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLPK 181

DB 150 VICQEQDNVRSFGLHFTYSVARRHPFLYAPAILSFVDFEHALQSCCKESDVGACLDTK 209

QY 182 LDELEDEGKASAKORLKCASIQRGFRFAKAVARLSORFPKAEVSKLAVTLDPKV 241

DB 210 EIVMEKAGKSVKQYFCGLTKQFGRVFOARQLIYLSQKYPKAPFSEVSKFVDSIGV 269

QY 242 HTECGHLLLECADRADLAKYICENQDSISKKKECEKPLLEKSHCIAEVENDEMPAD 301

DB 270 HKECEGDMVECDWARMNLCSSQDVFSGKIKDCEKPIVERSQSIMEAEFDEKPAD 329

QY 302 LPSLAADFVESKDCNKYAEAKQVFLGFMFLYEARRHPDYSVLLLLRAKTYETLEKCC 361

|    |     |           |                  |         |            |         |          |              |             |
|----|-----|-----------|------------------|---------|------------|---------|----------|--------------|-------------|
| Qy | 421 | PTLVEVRN  | LGKVGSKCKHPEAKMP | CAEDYL  | LSWLNQLCVL | LHEKTPV | SDRVTK   | CTCTES       | 480         |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |
| Db | 445 | PTLVEAARN | LRGVGKCKTFL      | EAQR    | LPVEDYL    | SAILNR  | LCVL     | LHEKTPVSEKVT | CKCSGS 504  |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |
| Qy | 481 | LVNRPCPS  | ALVDEVTYYPKFE    | NAETFF  | THADICTL   | SEKERO  | KKOTAL   | VELVHKHPKAT  | 540         |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |
| Db | 505 | LVYRPCPS  | ALVDEVTYYPKFE    | KATFTFF | THSDICTL   | LPDK    | QKIKQTAL | AELVHKHPKAT  | 564         |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |
| Qy | 541 | KEQLKAV   | MDFFAA           | FVEKCK  | ADDEKTC    | PAE     | SGKLV    | AAASQAAL     | 583         |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |
| Db | 565 | EDOLKTV   | MGDF             | FAQVD   | YDKCK      | KAADK   | DNCPAT   | SGPNLV       | ARSKEAL 607 |
|    |     | :         | :                | :       | :          | :       | :        | :            |             |

RESULT 8  
ABPSG

serum albumin precursor - pig (fragment)  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C/Accession: S01382; A61006  
R/Weinstock, J.; Baldwin, G.S.  
Nucleic Acids Res. 16, 9045, 1988  
A>Title: Nucleotide sequence of porcine liver albumin.  
A/Reference number: S01382; MUID:89016582; PMID:3174440  
A/Accession: S01382  
A/Status: translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-605 <WEB1>  
A/Cross-references: EMBL:XI2422; NID:g1875; PIDN:CAA30970.1; PID:g833798  
R/Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.  
J. Bone Miner. Res. 4, 235-241, 1989  
A>Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral  
A/Reference number: A61006; MUID:89265769; PMID:2728927  
A/Accession: A61006  
A/Molecule type: Protein  
A/Residues: 23-51, X, 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>  
A/Experimental source: dental enamel  
A/Note: albumin and other serum proteins are also found in bone  
C/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
teroid hormones (weak bonds with these hormones promote their transfer across the membra  
C/Superfamily: serum albumin; serum albumin repeat homology  
F,1-16/Domain: carrier protein; duplication; metal binding; plasma  
F,17-22/Domain: signal sequence (fragment) #status predicted <SIG>  
F,23-605/Product: serum albumin #status predicted <PRO>  
F,27-199/Domain: serum albumin #status predicted <MAT>  
F,218-391/Domain: serum albumin repeat homology <SA1>  
F,410-589/Domain: serum albumin repeat homology <SA2>  
F,75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4  
F,261/Binding site: bilirubin (lys) #status predicted

|                       |        |                      |  |            |     |        |      |
|-----------------------|--------|----------------------|--|------------|-----|--------|------|
| Query Match           | 77.7%; | Score                | 2411.5;                                      | DB         | 1;  | Length | 605; |
| Best Local Similarity | 76.0%; | Pred. No.            | 5e-152;                                      |            |     |        |      |
| Matches               | 438;   | Conservative         | 67;  | Mismatches | 70; | Indels | 1;   |
| Gaps                  | 1;     |                      |  |            |     |        |      |
| QY                    | 1      | DAHKSEVARRKDLGSENFKA | VLVLAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE      | 60         |     |        |      |
| DB                    | 23     | DTYKSEIARRKDLGEGYFKG | VLVLAFAQYLQCCPEEHHVKLVREVTETFAKTCVADESAE     | 82         |     |        |      |
| QY                    | 61     | NCDKSIHTLFGDKLCTVAT  | RTYEGEMADCCAKQSEPERNECFLOHKDNDPNPLRLVRPVP    | 120        |     |        |      |
| DB                    | 83     | NCDKSIHTLFGDKLCAIP   | SLREHYGLDACCKEEPEERNECFLOHKDNDPDKL-KPDE      | 141        |     |        |      |
| QY                    | 121    | DVWCYAFHNDNETFLKYL   | YEARRHPYFYPAPPELLFFPAKRYKAAFTCCCAADKKAACLLP  | 180        |     |        |      |
| DB                    | 142    | VALCADFQDEQKFWGKYL   | YEARRHPYFYPAPPELLYAIITYKDVFSCECCCAADKKAACLLP | 201        |     |        |      |
| QY                    | 181    | KLDELDEGKASAKORLKA   | SLQKFGERAFKAWAVARLSQRFPKFAEVSKLVTDLTK        | 240        |     |        |      |
| DB                    | 202    | KIEHLREKVLTSAAKORL   | KCAIQKSEGERAFKAWLSQRFPKADFTEISKIVTDLAK       | 261        |     |        |      |
| QY                    | 241    | VHTECCGGLLECADRADL   | AKYICENQDISISKLKECEKPLLEKSHCIAEVENDDEMA      | 300        |     |        |      |

Db 262 VHKECCHGDLLECADRADLAKYICENQDPTSTKLKECCDKPLLEKSHCIAEAKRDELPA 321  
 Qy 301 DLPSLAADFVESKDVCNKYAAEKDVLFGMLFYIYARRHPDYSVLLLRLLAKTETTTLKCC 360  
 Db 322 DLNPLEHDFVEDKEVCNKYKEAKDVLFTLYEYSRRHPDYSVLLLRLLAKIYEATLEDC 381  
 Qy 361 CAAADPHCEYAKVDFDEPKPLVERPONLIKNCELFEOLGEYKFONALLVRYTKVPOVST 420  
 Db 382 CAKEDPPACVATVDFKQPLVDFPPNLIKNCELFEKLGEYGFONALVRYTKVPOVST 441  
 Qy 421 PTLVEVSRLNGYSGKCKCKPEAKRMPCABEDYLSVLNQILCVLHEKTFVSDRVTKCCTES 480  
 Db 442 PTLVEVARKLGLVSGRCKPEBERLSCAEDYLSVLNRLCVLHEKTFVSEKVTCKCTES 501  
 Qy 481 LVNRRPCFSALEVDETVYVKEEFAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540  
 Db 502 LVNRRPCFSALTPTDETKPEFVEGTTFTFHADICTLPEDEKQIKKQATALLKHKHPAT 561  
 Qy 541 KEQLKAYMDDPFAAFVEKCKCKADDKETCFABEGKKLV 576  
 Db 562 BEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 597

RESULT 9  
 JC5838  
 alumin - Mongolian jird  
 C:Species: Meriones unguiculatus (Mongolian jird)  
 C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
 JC5838  
 C:Accession: JC5838  
 R:Yoshida, K.; Seito-Ohshima, A.; Sinohara, H.  
 DNA Res. 4, 351-354, 1997  
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthe  
 A:Reference number: JC5838; MUID:98116663; PMID:9455485  
 A:Accession: JC5838  
 A:Molecule type: mRNA  
 A:Residues: 1-609 <YOS>  
 A:Cross-references: DBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2317278  
 A:Experimental source: liver  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:222-395/Domain: serum albumin repeat homology <SA>

Query Match 76.9%; Score 2387; DB 2; Length 609;  
 Best Local Similarity 73.9%; Prid. No. 2.1e-150;  
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0

Qy 2 AHKSEVARFVDLGENSKALVLIFAQVLOCPEDHVKLVNEYTEFAKTCVADESAEN 61  
 Db 27 AHKSEIARHYKDLGEGYKFGLVLYTFSQLQCKSTEHHVKLVREVTDFASNCAKESAEN 86  
 Qy 62 CDKSLHTLFGDKLCITVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEVD 121  
 Db 87 CDKSLHTLFGDKLCSLPNGEXYEMADCCAKQEPERNECFLOHKDDNPNLPPFKAEED 146  
 Qy 122 VMCTAFHDNEETFLKKYLIEARRHPYFVAPPELLFPKRYKAAFTCCOAAADKAACLLPK 181  
 Db 147 AMCTAFQENAEAFMGHYLHEVARRRHPYFVGPPELLVLAATYAVLTCECAADKGACLTPK 206  
 Qy 182 LDELDEGKASAKORLLKASLQKFGERAFKAWAVARLUSQRPPKAEFAEVSKLVTDLTKV 241  
 Db 207 LDALKKALVSAVRQLKCSSMKTKGERAFKAWAVARMSQTFPNADFAEITKLATDLTKV 266  
 Qy 242 HTECCHGDLLECADRADLAKYICENQDSISKKLKECCEKPLLEKSHCIAEAVENDMPAD 301  
 Db 267 TOECCHGDLLECADRAELAKYCNENQASISKKQACCKEMLOKSCQCLAEVEHDDMPAD 326  
 Qy 302 LPSLAADFVESKDVCNKYAAEKDVLFGMLFYIYARRHPDYSVLLLRLLAKTETTTLKCC 361  
 Db 327 LPALTADFVEDKDVCNKYAAEKDVLFTLYEYSRRHPYVSLLLRLLAKKYEATLEKCC 386  
 Qy 362 AAADPHCEYAKVDFDEPKPLVERPONLIKNCELFEOLGEYKFONALLVRYTKVPOVST 421  
 Db 387 AEADPHACYGHVDFDEPKPLVEBPQNIAVNGNCLEYKLGEYGFONALVRYTKVPOVST 446

Best Local Similarity 75.0%; Pred. No. 28-153;  
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHKEVAHFKDGLBENFKALVIAFAQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 60  
DB 25 DTHKSTIAHFNGLGEENFGVLIAFSYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
DB 85 GCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 143  
QY 121 DVMCTAFHNEETFLKKYLYEARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
DB 144 DTLCAEFKADKFKWGLYEVARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 203  
QY 181 KIDELDEGKASSAKORLKASIQKFGERAFKAWAVARLSQRPKXAEPAEVSGLVDTLTK 240  
DB 204 KIDAMEKVLASSARQLRCASIQKFGERALKAWSVARLSQRPKADFTDVTKIVTDLTK 263  
QY 241 VHTCECHGDLLECCADRADLAKYICENODSISKLKCECKPILKSKHICIAEVNDMPA 300  
DB 264 VHKCECHGDLLECCADRADLAKYICDHQDALSSKLKCECDKPVLEKSHICIAEVNDMPA 323  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 324 NLPLTADFAEDKEVKCKYQAEKDVFLGSLFYESSRRHPYAVSVLLRLAKEYEATLEDC 383  
QY 361 CAADPHCYAKVDFEKFPLVSEPNLIIKONCELFQOLGEYKFNALLVRYTKYQVST 420  
DB 384 CAKEDPHACYATVDFKLKHLVDEPNLIIKONCELFKGEYGFONALLVRYTKYQVST 443  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCTES 480  
DB 444 PTLVEISRLNGKVGSKCKHPEAKRMPCEDEYLSVLNOLCVLHEKTPVSEKVKCTES 503  
QY 481 LVNRRPFCFSALEVDVETVPKFNATFTTHADICTLSEKERQIKKOTALVELVHKPKAT 540  
DB 504 LVNRRPFCFSDLTLDVETVPKFPDEKFTTHADICTLPTDEKQIKKOTALVELVHKPKAT 563  
QY 541 KEQLKAWMDFAFVEKCKKADDEKCTFAEKGKLVAAASQAL 583  
DB 564 DEQKUTVNEFVAFVDCCKADDEKCGFVLEGKLVASTQAL 606

## RESULT 7

ABRTS  
serum albumin precursor - rat  
N;Alcinate names: preproalbumin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999  
C;Accession: A93872; A92211; A91946; A91940; C45800; 157621; A03233  
R;Satgent, T.D.; Yang, M.; Bonner, J.  
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981  
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.  
A;Reference number: A93872; MUID:81223722; PMID:7017712  
A;Accession: A93872  
A;Molecule type: mRNA  
A;Residues: 1-608 <SAR>  
A;Cross-references: GB:J00698; NID:G55627; PID:CAA24532.1; PID:G55628  
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.  
J. Biol. Chem. 252, 6846-6855, 1977  
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis  
A;Reference number: A92211; MUID:77243657; PMID:893447  
A;Note: cleavages during protein maturation  
A;Accession: A92211  
A;Molecule type: protein  
A;Residues: 1-38 <STR>  
R;Isemura, S.; Ikenaka, T.  
J. Biochem. 83, 35-48, 1978  
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage  
A;Reference number: A91946; MUID:78109429; PMID:564345  
A;Accession: A91946  
A;Molecule type: protein

A;Residues: 25-222 <ISI>  
R;Isemura, S.; Ikenaka, T.  
J. Biochem. 79, 1183-1196, 1976  
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino  
A;Reference number: A91940; MUID:76260153; PMID:956149  
A;Accession: A91940  
A;Molecule type: protein  
A;Residues: 223-288/572-608 <IS2>  
A;Note: 262-Leu was also found  
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A;Reference number: A90758; MUID:79001617; PMID:80265  
A;Contents: annotation; copper binding  
R;Caraway, R.E.; Cochran, D.G.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid protease  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: C45800  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 166-173 <CAR>  
R;Heard, J.  
Mol. Cell. Biol. 7, 2425-2434, 1987  
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improve  
A;Reference number: 157621; MUID:87286876; PMID:3475566  
A;Accession: 157621  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-5 <RES>  
A;Cross-references: GB:M16925; NID:G202828; PID:AAA40712.1; PID:G554412  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-608/Product: serum albumin #status experimental <MAT>  
F;29-202/Domain: serum albumin repeat homology <SAL>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;77/Binding site: copper (His) #status experimental  
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393.

Query Match 78.2%; Score 2426; DB 1; Length 608;  
Best Local Similarity 73.4%; Pred. No. 5.5e-153;  
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHKEVAHFKDGLBENFKALVIAFAQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 60  
DB 25 EAHKSEIAHFRKDLGEOHFKGLVLIAPSYLOQCPYEEHKLKQVETDFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 144  
QY 121 DVMCTAFHNEETFLKKYLYEARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
DB 145 EAMCTSFQENPTSFGLHYLHEVARRHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 204  
QY 181 KIDELDEGKASSAKORLKASIQKFGERAFKAWAVARLSQRPKXAEPAEVSGLVDTLTK 240  
DB 205 KIDAVEKALVAARVQRWCKSSNQFGERAFKAWAVARLSQRPKXAEPAEVSGLVDTLTK 264  
QY 241 VHTCECHGDLLECCADRADLAKYICENODSISKLKCECKPILKSKHICIAEVNDMPA 300  
DB 265 INKECHGDLLECCADRADLAKYICENODSISKLKCECKPILKSKHICIAEVNDMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 384  
QY 361 CAADPHCYAKVDFEKFPLVSEPNLIIKONCELFQOLGEYKFNALLVRYTKYQVST 420  
DB 385 CAGDPPACYGTVLAEFQPLVEEPKLVKNTCNCELYKLGEGYGFONALLVRYTKYQVST 444

Query Match 78.8%; Score 2446.5; DB 1; Length 607;

Query Match 84.4%; Score 2620; DB 2; Length 608;  
 Best Local Similarity 82.0%; Pred. No. 8.1e-166;  
 Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHRKDLGSENFKALVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 25 BAHQSEIAHFRNDLGEHFRGLVAFVAFQYLOOCPFFEDHVKLVNEVTEFAKTCVADSEAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120  
 DB 85 NCKSLHLLGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 144

QY 121 DVNCTAFHNEBTEFLKXLYEYARHPYFAPPELLFAKRYKAAFTCCOAAADKAAACLLP 180  
 DB 145 DAMCTAFHNEBTEFLKXLYEYARHPYFAPPELLFAKRYKAAFTCCOAAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYKLVITDLTK 240  
 DB 205 KYDALREKVLASSAKERLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYKLVITDLTK 264

QY 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 DB 265 IHKECCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324

QY 301 DPLSLAADFVSKDCKNVAEAKDVFGLMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360  
 DB 325 DPLSLAADFVSKDCKNVAEAKDVFGLMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384

QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALVRYTKKVPQVST 420  
 DB 385 CATDDPPACYAHVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALVRYTKKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHKTVPVSDRVTKCCTES 480  
 DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHKTVPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTPHADICTLSEKQIKQTALVELVKKHPRAT 540  
 DB 505 LVNRRPCFSALEVDETYVPKFEFNAETFTPHADICTLSEKQIKQTALVELVKKHPRAT 564

QY 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAASQAL 583  
 DB 565 BEQLKTVMGDFGSDVFDKCAAEDEKACFAEFGKLVAAASQAL 607

RESULT 4  
 ABBOS  
 serum albumin precursor - horse  
 C/Species: Equus caballus (domestic horse)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C/Accession: S34053  
 R/Ho, J. X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
 Eur. J. Biochem. 215, 205-212, 1993  
 A/Title: X-ray and Primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
 A/Reference number: S34053; MUID:93345495; PMID:8344282  
 A/Accession: S34053  
 A/Molecule type: mRNA  
 A/Residues: 1-607 <HOA>  
 A/Cross-references: GB:X74045; NID:G399671; PID:CAA52194.1; PID:G399672  
 C/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)  
 C/Superfamily: serum albumin; serum albumin repeat homology  
 C/Keywords: carrier protein; duplication; metal binding; plasma  
 F/1-18/Domain: signal sequence #status predicted <SIG>  
 F/19-24/Domain: propeptide #status predicted <PRO>  
 F/25-607/Product: serum albumin #status predicted <MAT>  
 F/25-201/Domain: serum albumin repeat homology <SA1>  
 F/220-393/Domain: serum albumin repeat homology <SA2>  
 F/412-391/Domain: serum albumin repeat homology <SA3>  
 F/27/Binding site: copper (His) #status predicted  
 F/77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4  
 F/263/Binding site: bilirubin (Lys) #status predicted

Query Match 79.8%; Score 2475.5; DB 1; Length 607;  
 Best Local Similarity 76.3%; Pred. No. 2.9e-156;  
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHSEVAHRKDLGSENFKALVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 25 DTHKSTIAHFRNDLGEHFRGLVAFVAFQYLOOCPFFEDHVKLVNEVTEFAKTCVADSEAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 143

QY 121 DVNCTAFHNEBTEFLKXLYEYARHPYFAPPELLFAKRYKAAFTCCOAAADKAAACLLP 180  
 DB 144 DAOCAAFQDDEPKFLGKLYEYARHPYFAPPELLFAKRYKAAFTCCOAAADKAAACLLP 203

QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYKLVITDLTK 240  
 DB 204 KLDALKERILLSSAKERLKCSFQNGERAFKAWAVARLSORFPKAEFAEYKLVITDLTK 263

QY 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 DB 264 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 323

QY 301 DPLSLAADFVSKDCKNVAEAKDVFGLMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360  
 DB 324 DPLSLAADFVSKDCKNVAEAKDVFGLMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 383

QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALVRYTKKVPQVST 420  
 DB 384 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALVRYTKKVPQVST 443

QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHKTVPVSDRVTKCCTES 480  
 DB 444 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHKTVPVSDRVTKCCTES 503

QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTPHADICTLSEKQIKQTALVELVKKHPRAT 540  
 DB 504 LABRRPCFSALEVDETYVPKFEFNAETFTPHADICTLSEKQIKQTALVELVKKHPRAT 563

QY 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAASQAL 583  
 DB 564 KEQLKTVMGDFGSDVFDKCAAEDEKACFAEFGKLVAAASQAL 606

RESULT 5  
 ABBOS  
 serum albumin precursor [validated] - bovine  
 N/Alternate names: 67K protein; preproalbumin  
 C/Species: Bos primigenius taurus (cattle)  
 C/Date: 24-Apr-1984 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000  
 C/Accession: A38885; A36401; A31258; B60808; S10780; D45800; A28693; A90309; A91458; A9  
 R/Holowachuk, E.W.; Stolttenborg, J.K.; Reed, R.G.; Peters Jr., T.  
 submitted to the EMBL Data Library, August 1991  
 A/Description: Bovine serum albumin: cDNA sequence and expression.  
 A/Reference number: A38885  
 A/Accession: A38885  
 A/Molecule type: mRNA  
 A/Residues: 1-607 <HOL>  
 A/Cross-references: EMBL:M73215  
 R/Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
 Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
 A/Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
 A/Reference number: A36401; MUID:91083649; PMID:2260975  
 A/Accession: A36401  
 A/Molecule type: protein  
 A/Residues: 25-41, 'H', '43-189, 'E', '91-213, 'T', '215-323, 'D', '325-393, 'TS', '396-607 <HTR>  
 R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
 Eur. J. Biochem. 98, 477-485, 1979  
 A/Title: Bioynthesis of bovine plasma proteins in a cell-free system.  
 A/Reference number: A31258; MUID:80024278; PMID:488109  
 A/Accession: A31258  
 A/Molecule type: protein

F:166-174/Product: kinetensin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:27/Binding site: copper (His) #status predicted  
 F:77-86/99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,413-592/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 9e-198;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFDLGEENFKALVLIAPAYIQOOCFFEDHVKLVNVEVTEFAKTCVADSEAE 60  
 DB 25 DAHKSEVAHRFDLGEENFKALVLIAPAYIQOOCFFEDHVKLVNVEVTEFAKTCVADSEAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144  
 QY 121 DVMCTAFHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
 DB 145 DVMCTAFHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204  
 QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAPFAEVSCLVTDLT 240  
 DB 205 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAPFAEVSCLVTDLT 264  
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
 DB 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKDVCKNYABAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 360  
 DB 325 DLPSLAADFVESKDVCKNYABAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 384  
 QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
 DB 385 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 444  
 QY 421 PTLVEVSRLNKGKVGKCKCHPEAKMPCAEYDLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEVSRLNKGKVGKCKCHPEAKMPCAEYDLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540  
 DB 505 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 564  
 QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585  
 DB 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 609

RESULT 2  
 A47391  
 serum albumin precursor - rhesus macaque  
 C:Species: Macaca mulatta (rhesus macaque)  
 C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
 C:Accession: A47391  
 R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding  
 A:Reference number: A47391; MUID:93211971; PMID:8460152  
 A:Contents: B/B homozygote  
 A:Accession: A47391  
 A:Status: preliminary  
 A:Molecule type: mRNA; protein  
 A:Residues: 1-600 <WAT>  
 A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295  
 A:Experimental source: liver  
 A:Note: Sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281)  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:21-194/Domain: serum albumin repeat homology <SA1>  
 F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2942; DB 2; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 4e-187;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFDLGEENFKALVLIAPAYIQOOCFFEDHVKLVNVEVTEFAKTCVADSEAE 60  
 DB 17 DTHKSEVAHRFDLGEHFFKGLVLAFAQYLOOCFFEEHVKLNVNEVTEFAKTCVADSEAE 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 136  
 QY 121 DVMCTAFHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
 DB 137 DVMCTAFHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 196  
 QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAPFAEVSCLVTDLT 240  
 DB 197 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAPFAEVSCLVTDLT 256  
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
 DB 257 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 316  
 QY 301 DLPSLAADFVESKDVCKNYABAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 360  
 DB 317 DLPSLAADFVESKDVCKNYABAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 376  
 QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
 DB 377 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 436  
 QY 421 PTLVEVSRLNKGKVGKCKCHPEAKMPCAEYDLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 437 PTLVEVSRLNKGKVGKCKCHPEAKMPCAEYDLSVVLNQLCVLHEKTPVSEKVTCKCTES 496  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540  
 DB 497 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 556  
 QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKGLVAASQAAL 583  
 DB 557 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKGLVAASQAAL 599

RESULT 3  
 S57632  
 serum albumin precursor - cat  
 C:Species: Felis silvestris catus (domestic cat)  
 C>Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: JC4660; S57632  
 R:Hilger, C.; Grigioni, F.; Hentges, F.  
 Gene 169, 295-296, 1996  
 A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
 A:Reference number: JC4660; MUID:96194824; PMID:8647469  
 A:Accession: JC4660  
 A:Molecule type: mRNA  
 A:Residues: 1-608 <HI2>  
 A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
 A:Experimental source: liver  
 C:Comment: This protein is the major protein component in plasma. It functions as a multimeric protein with 35 conserved cysteine residues.  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: liver; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status predicted <PRP>  
 F:25-608/Product: serum albumin #status predicted <MAT>  
 F:221-394/Domain: serum albumin repeat homology <SA1>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
R;Menya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A;Reference number: G08292  
A;Accession: G01747  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-120, 'G', 122-455 <MEN>  
A;Cross-references: EMBL:U22961; NID:G763428; PID:G763431  
P;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
A;Reference number: S55314; MUID:95275251; PMID:7755581  
A;Accession: S55314  
A;Molecule type: protein  
A;Residues: 19-27 <LED>  
P;Meloun, B.; Moravsek, L.; Kostka, V.  
FEBS Lett. 56, 134-137, 1975  
A;Title: Complete amino acid sequence of human serum albumin.  
A;Reference number: A91420; MUID:76187907; PMID:1225573  
A;Accession: A91420  
A;Molecule type: protein  
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
P;Roehr, U.; Spittler, G.; Tripler, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A;Reference number: S06422  
A;Note: this paper is in German, with an English abstract  
A;Accession: S06422  
A;Molecule type: protein  
A;Residues: 25-48 <ROE>  
P;Finch, J.W.; Crouch, R.K.; Knaap, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A;Title: Mass spectrometric identification of modifications to human serum albumin treat  
A;Reference number: S36882; MUID:93384321; PMID:8737198  
A;Accession: S36882  
A;Molecule type: protein  
A;Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>  
P;Kausler, E.; Spittler, G.  
Biochem. Biophys. Res. Commun. 136, 849-855, 1991  
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol  
A;Reference number: S17599; MUID:92126241; PMID:1772598  
A;Accession: S17599  
A;Molecule type: protein  
A;Residues: 25-54, 354-357, 431-447 <KAU>  
A;Note: 49-Leu was also found  
P;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1580-1584, 1989  
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A;Reference number: A45800; MUID:89341406; PMID:2474609  
A;Accession: A45800  
A;Molecule type: protein  
A;Residues: 166-173, 'L', <MOG>  
P;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
Biochem. Biophys. Res. Commun. 136, 993-998, 1986  
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
A;Reference number: A03239; MUID:86242180; PMID:3087352  
A;Accession: A03239  
A;Molecule type: protein  
A;Residues: 166-173, 'L', <MOG>  
P;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A;Title: Mutations in genetic variants of human serum albumin found in Italy.  
A;Reference number: A38255; MUID:91062352; PMID:2247440  
A;Accession: A38255  
A;Molecule type: protein  
A;Residues: 76-111 <GAL1>  
A;Accession: B38255  
A;Molecule type: protein  
A;Residues: 82-105, 'K', 107-110 <GAL2>  
A;Note: this variant is designated albumin Vibo Valentia  
A;Accession: A38255  
A;Molecule type: protein

A;Residues: 76-83, 'K', 85-106 <GAL3>  
A;Note: this variant is designated albumin Torino  
P;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A;Title: The structural characterization and bilirubin-binding properties of albumin He  
A;Reference number: S32298; MUID:93292504; PMID:8513793  
A;Accession: S32298  
A;Molecule type: protein  
A;Residues: 255-263, 'E', 265-281 <MIN1>  
A;Note: this variant is designated albumin Herborn  
P;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe  
A;Reference number: S21078; MUID:92190239; PMID:1347703  
A;Accession: S21078  
A;Molecule type: protein  
A;Residues: 354-356, 'K', 358-378 <MIN2>  
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported  
P;He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A;Title: Atomic structure and chemistry of human serum albumin.  
A;Reference number: A46756; MUID:92334427; PMID:1630489  
A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
P;Brown, J.R.; Shockley, P.; Behrens, P.Q.  
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40  
A;Reference number: A94442  
A;Contents: annotation; three-dimensional structure and disulfide bonds  
P;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A;Title: Disulfide bonds in human serum albumin.  
A;Reference number: A90930  
A;Contents: annotation; disulfide bonds  
P;Jacobsen, C.  
Biochem. J. 171, 453-459, 1978  
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
A;Reference number: A90239; MUID:78186630; PMID:656055  
A;Contents: annotation; bilirubin-binding site  
P;Peters, T.; Reed, R.G.  
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20  
A;Title: Serum albumin: conformation and active sites.  
A;Reference number: A94408  
A;Contents: annotation; binding sites  
P;Harper, M.E.; Dugaiczky, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene  
A;Reference number: A90028; MUID:8327982; PMID:6192711  
A;Contents: annotation; gene position  
P;Walker, J.S.  
FEBS Lett. 66, 173-175, 1976  
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid  
A;Reference number: A46755; MUID:76257808; PMID:955075  
A;Contents: annotation  
P;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosph  
A;Reference number: A56234; MUID:92183981; PMID:1544460  
A;Contents: annotation  
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in  
case activity  
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
C;Comment: A large number of variants of human serum albumin have been described.  
C;Genetics:  
A;Gene: GDB:ALB  
A;Cross-references: GDB:118990; OMIM:103600  
A;Map position: 4q11-4q13  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid  
P;1-18/Domain: signal sequence; #status predicted <SIG>  
P;19-24/Domain: propeptide; #status experimental <PRO>  
P;25-609/Product: serum albumin repeat homology <MPT>  
P;25-202/Domain: serum albumin repeat homology <SA1>



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 71.3019 Seconds  
(without alignments)  
789.208 Million cell updates/sec

Title: US-09-832-929-18  
Perfect score: 3103  
Sequence: 1 DAHSEVAHRFDLGEENFK.....TCFAEKGKLVASQAALGL 595

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR 78:\*

1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | ID       | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 3103   | 100.0       | 609    | 1 ABHUS  | serum albumin prec |
| 2          | 2942   | 94.8        | 600    | 2 A47391 | serum albumin prec |
| 3          | 2620   | 84.4        | 608    | 2 S57632 | serum albumin prec |
| 4          | 2475.5 | 79.8        | 607    | 1 ABHOS  | serum albumin prec |
| 5          | 2446.5 | 78.8        | 607    | 1 ABBSH  | serum albumin prec |
| 6          | 2432.5 | 78.4        | 607    | 1 ABBSH  | serum albumin prec |
| 7          | 2426   | 78.2        | 608    | 1 ABRTS  | serum albumin prec |
| 8          | 2411.5 | 77.7        | 605    | 1 ABPGS  | serum albumin prec |
| 9          | 2387   | 76.9        | 609    | 2 JCS938 | albumin - Mongolia |
| 10         | 1861   | 60.0        | 453    | 2 A05139 | serum albumin - mo |
| 11         | 1557.5 | 50.2        | 615    | 1 ABCHS  | serum albumin prec |
| 12         | 1253.5 | 40.4        | 609    | 2 JCA258 | alpha-fetoprotein  |
| 13         | 1249.5 | 40.3        | 609    | 1 PPHU   | alpha-fetoprotein  |
| 14         | 1242.5 | 40.0        | 609    | 1 PFCO   | alpha-fetoprotein  |
| 15         | 1205   | 38.8        | 607    | 1 ABXL72 | 74K albumin prec   |
| 16         | 1181.5 | 38.1        | 265    | 2 I46986 | albumin - dog (fra |
| 17         | 1175.5 | 37.9        | 608    | 1 ABXL68 | 68K serum albumin  |
| 18         | 1084   | 34.9        | 605    | 1 FPMS   | alpha-fetoprotein  |
| 19         | 1057   | 34.4        | 611    | 1 FPRT   | alpha-fetoprotein  |
| 20         | 1055   | 34.0        | 599    | 1 A54906 | afanin precursor - |
| 21         | 928.5  | 29.9        | 614    | 2 S59517 | serum albumin prec |
| 22         | 928    | 29.9        | 608    | 2 A53195 | afanin precursor - |
| 23         | 747.5  | 24.1        | 608    | 1 ABONS1 | serum albumin 1 pr |
| 24         | 742.5  | 23.9        | 608    | 1 ABONS2 | serum albumin 2 pr |
| 25         | 699    | 22.5        | 382    | 2 A37253 | serum albumin - bu |
| 26         | 440.5  | 14.2        | 1423   | 1 S27941 | serum albumin - se |
| 27         | 386    | 12.4        | 474    | 1 VYHUD  | vitamin D-binding  |
| 28         | 385    | 12.4        | 476    | 1 VYRUD  | vitamin D-binding  |
| 29         | 372    | 12.0        | 472    | 1 A35327 | vitamin D-binding  |

## ALIGNMENTS

serum albumin precursor [validated] - human  
N;Alternate names: preproalbumin

N;Contains: kinetensin

C;Species: Homo sapiens (man)

C;Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text, change 17-Mar-2000

C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3

R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu

Nucleic Acids Res. 9, 6103-6114, 1981

A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col

A;Reference number: A93743; MUID:82081882; PMID:6171778

A;Accession: A93743

A;Molecule type: mRNA

A;Residues: 1-419, 'K', 421-609 <LAW>

A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:CAA2

R;Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936

A;Molecule type: mRNA

A;Residues: 1-120, 'G', 122-609 <DUG>

A;Cross-references: EMBL:V00494; NID:g28589; PIDN:CAA23753.1; PID:g28590

R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and

A;Reference number: I39427; MUID:86140099; PMID:2419329

A;Accession: I39427

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <URA>

A;Cross-references: GB:M13075; NID:g179330; PIDN:AAAS1688.1; PID:g553173

R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family

A;Reference number: I59286; MUID:94181575; PMID:8134387

A;Accession: I59286

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 282-290, 'KSRFDLQ' <WAT>

A;Cross-references: GB:S69192; NID:g546032; PIDN:AAB30282.1; PID:g546033

A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl

A;Reference number: I59313; MUID:94294404; PMID:8022807

A;Accession: I59313

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 589-590, 'ALPVRVKNLLQVQLP' <MAD>

A;Cross-references: GB:S70799; NID:g547231; PIDN:AAB31177.1; PID:g547232



```
QY 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

RESULT 14
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 2068; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
Db 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRE 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRE 120
QY 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
```

```
RESULT 15
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Iau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
```

```
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1

Query Match 100.0%; Score 2068; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
Db 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRE 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRE 120
QY 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKVCKNYAEAKDVFGLGFLYFYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

Search completed: April 19, 2004, 12:54:57
Job time : 163.947 secs
```

```

; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCNKYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCNKYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAADPHECYAKVDFEFPKPLVEBPQNL 387
DB 361 CAADPHECYAKVDFEFPKPLVEBPQNL 387

RESULT 13
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 2002-05-24
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 2068; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300

```

```

; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300

```

QY 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
QY 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 9

US-09-833-245-18  
; Sequence 18, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-245-18

Query Match 100.0%; Score 2068; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
QY 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 10

US-10-424-999-11  
; Sequence 11, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 11  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Fusion protein human abrogen  
US-10-424-999-11

Query Match 100.0%; Score 2068; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240  
QY 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLCADDADRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 11

US-10-425-000-31  
; Sequence 31, Application US/10425000  
; Publication No. US20040052777A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit

; APPLICANT: Beltzer, James P.  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Fleming, Tony J.  
; APPLICANT: Ladner, Robert Charles  
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)  
; FILE REFERENCE: DXX-018.1 PCT; DXX-018.1 US  
; CURRENT APPLICATION NUMBER: US/09/932,322  
; CURRENT FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 445  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-932-322-445

Query Match 100.0%; Score 2068; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Qy 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360  
Qy 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

## RESULT 7

US-09-832-501-18  
; Sequence 18, Application US/09832501  
; Publication No. US20030199043A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J.  
; APPLICANT: Sleep, Darrell  
; APPLICANT: Turner, Andrew J.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF542  
; CURRENT APPLICATION NUMBER: US/09/832,501  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-832-501-18

Query Match 100.0%; Score 2068; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Qy 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360  
Qy 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

## RESULT 8

US-09-833-118-18  
; Sequence 18, Application US/09833118  
; Publication No. US20030219875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF544  
; CURRENT APPLICATION NUMBER: US/09/833,118  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-118-18

Query Match 100.0%; Score 2068; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 1 DAHSEVAHFRKDLGKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Qy 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQKPERNECFLOHKKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 4

US-09-833-041-18  
; Sequence 18, Application US/09833041  
; Publication No. US20030125247A1

## GENERAL INFORMATION:

; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF545  
; CURRENT APPLICATION NUMBER: US/09/833,041  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens

US-09-833-041-18

Query Match 100.0%; Score 2068; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190; Mismatches 0; Indels 0; Gaps 0;  
Matches 387; Conservative 0;

QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360

DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 5

US-09-833-117-18  
; Sequence 18, Application US/09833117  
; Publication No. US20030171267A1

## GENERAL INFORMATION:

; APPLICANT: Rosen, Craig A.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF543  
; CURRENT APPLICATION NUMBER: US/09/833,117  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens

US-09-833-117-18

Query Match 100.0%; Score 2068; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.3e-190; Mismatches 0; Indels 0; Gaps 0;  
Matches 387; Conservative 0;

QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

## RESULT 6

US-09-932-322-445

; Sequence 445, Application US/09932322  
; Publication No. US20030194743A1  
; GENERAL INFORMATION:  
; APPLICANT: Dyax Corp.

```
;
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 2068; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLTK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLTK 360
QY 361 CAADPHECYAKVDFEKPILVEBPQNL 387
DB 361 CAADPHECYAKVDFEKPILVEBPQNL 387

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match      100.0%; Score 2069; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 162.947 Seconds  
(without alignments)  
654.724 Million cell updates/sec

Title: US-09-832-929-18\_COPY\_1\_387

Perfect score: 2068

Sequence: 1 DAKSEVAHRFKDLGEENFK.....ECYAKVFDEPKPLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/FCI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/FCIUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 2068  | 100.0       | 585    | 9  | US-09-929-552-2   |
| 2          | 2068  | 100.0       | 585    | 10 | US-09-932-613-445 |
| 3          | 2068  | 100.0       | 585    | 10 | US-09-984-010-26  |
| 4          | 2068  | 100.0       | 585    | 10 | US-09-833-041-18  |
| 5          | 2068  | 100.0       | 585    | 10 | US-09-833-117-18  |
| 6          | 2068  | 100.0       | 585    | 10 | US-09-932-322-445 |
| 7          | 2068  | 100.0       | 585    | 10 | US-09-832-501-18  |
| 8          | 2068  | 100.0       | 585    | 11 | US-09-833-118-18  |
| 9          | 2068  | 100.0       | 585    | 11 | US-09-833-245-18  |
| 10         | 2068  | 100.0       | 585    | 12 | US-10-424-999-11  |
| 11         | 2068  | 100.0       | 585    | 12 | US-10-425-000-31  |
| 12         | 2068  | 100.0       | 585    | 12 | US-10-433-108-34  |
| 13         | 2068  | 100.0       | 585    | 13 | US-10-153-064-5   |
| 14         | 2068  | 100.0       | 585    | 14 | US-10-153-604A-5  |
| 15         | 2068  | 100.0       | 585    | 14 | US-10-319-263-1   |

|    |      |       |     |    |                    |                    |
|----|------|-------|-----|----|--------------------|--------------------|
| 16 | 2068 | 100.0 | 585 | 14 | US-10-319-263-2    | Sequence 2, Appli  |
| 17 | 2068 | 100.0 | 585 | 14 | US-10-414-469-1    | Sequence 1, Appli  |
| 18 | 2068 | 100.0 | 585 | 14 | US-10-414-469-2    | Sequence 2, Appli  |
| 19 | 2068 | 100.0 | 585 | 14 | US-10-413-831-1    | Sequence 1, Appli  |
| 20 | 2068 | 100.0 | 585 | 14 | US-10-413-831-2    | Sequence 2, Appli  |
| 21 | 2068 | 100.0 | 585 | 15 | US-10-413-832-1    | Sequence 1, Appli  |
| 22 | 2068 | 100.0 | 585 | 15 | US-10-413-832-2    | Sequence 2, Appli  |
| 23 | 2068 | 100.0 | 585 | 15 | US-10-414-386-1    | Sequence 1, Appli  |
| 24 | 2068 | 100.0 | 585 | 15 | US-10-414-386-2    | Sequence 2, Appli  |
| 25 | 2068 | 100.0 | 585 | 15 | US-10-233-675A-11  | Sequence 11, Appli |
| 26 | 2068 | 100.0 | 585 | 15 | US-10-462-262-26   | Sequence 26, Appli |
| 27 | 2068 | 100.0 | 604 | 10 | US-09-984-010-7    | Sequence 7, Appli  |
| 28 | 2068 | 100.0 | 609 | 10 | US-09-919-039-370  | Sequence 370, App  |
| 29 | 2068 | 100.0 | 609 | 12 | US-10-609-346-12   | Sequence 12, Appli |
| 30 | 2068 | 100.0 | 609 | 13 | US-10-153-064-7    | Sequence 7, Appli  |
| 31 | 2068 | 100.0 | 609 | 14 | US-10-153-604A-7   | Sequence 7, Appli  |
| 32 | 2068 | 100.0 | 609 | 14 | US-10-365-623-23   | Sequence 23, Appli |
| 33 | 2068 | 100.0 | 610 | 9  | US-09-984-186-2    | Sequence 2, Appli  |
| 34 | 2068 | 100.0 | 610 | 14 | US-10-237-687-2    | Sequence 2, Appli  |
| 35 | 2068 | 100.0 | 610 | 14 | US-10-237-708-2    | Sequence 2, Appli  |
| 36 | 2068 | 100.0 | 610 | 14 | US-10-237-866-2    | Sequence 2, Appli  |
| 37 | 2068 | 100.0 | 610 | 14 | US-10-237-871-2    | Sequence 2, Appli  |
| 38 | 2068 | 100.0 | 610 | 14 | US-10-237-624-2    | Sequence 2, Appli  |
| 39 | 2068 | 100.0 | 616 | 12 | US-10-433-108-13   | Sequence 13, Appli |
| 40 | 2068 | 100.0 | 624 | 12 | US-10-433-108-16   | Sequence 16, Appli |
| 41 | 2068 | 100.0 | 631 | 12 | US-10-433-108-14   | Sequence 14, Appli |
| 42 | 2068 | 100.0 | 640 | 12 | US-10-433-108-15   | Sequence 15, Appli |
| 43 | 2068 | 100.0 | 640 | 12 | US-10-433-108-17   | Sequence 17, Appli |
| 44 | 2068 | 100.0 | 651 | 13 | US-10-153-064-133  | Sequence 133, App  |
| 45 | 2068 | 100.0 | 651 | 14 | US-10-153-604A-133 | Sequence 133, App  |

#### ALIGNMENTS

#### RESULT 1

US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonnschein, Carlos  
; Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MERRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids